



D9/718.342

INVESTOR IN PEOPLE

The Patent Office  
Concept House  
Cardiff Road  
Newport  
South Wales  
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

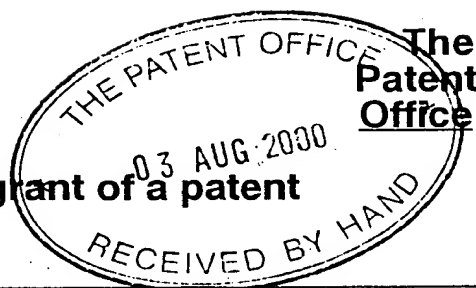


Signed

Dated: 28 November 2000

**THIS PAGE BLANK (USPTO)**

Patents Form 1/77  
Patents Act 1977  
(Rule 16)



04AUG00 E558354-1 D02917  
P01/7700 0.00-0019120.5

## Request for grant of a patent

The Patent Office  
Cardiff Road  
Newport  
Gwent NP9 1RH

1. Your reference  
2636002/DS

2. Patent Application Number  
**0019120.5**      03 AUG 2000

3. Full name, address and postcode of the or of each applicant (*underline all surnames*)

Canon Kabushiki Kaisha  
30-2 3-Chome Shimomaruko  
Ohta-Ku  
Tokyo  
Japan

36301 0003

Patents ADP number (*if known*)

If the applicant is a corporate body, give the  
country/state of its incorporation

Country: JAPAN  
State:

4. Title of the invention  
**IMAGE PROCESSING APPARATUS**

5. Name of agent  
"Address for Service" in the United Kingdom  
to which all correspondence should be sent  
Patents ADP number

Beresford & Co  
2/5 Warwick Court  
High Holborn  
London WC1R 5DJ

1826001

6. Priority details

Country      Priority application number      Date of filing

# Patents Form 1/77

7. If this application is divided or otherwise derived from an earlier UK application give details

Number of earlier application

Date of filing

8. Is a statement of inventorship and or right to grant of a patent required in support of this request?

YES

9. Enter the number of sheets for any of the following items you are filing with this form.

Continuation sheets of this form 0

Description 55

Claim(s) 19

Abstract 1

Drawing(s) 23

23 2

10. If you are also filing any of the following, state how many against each item.

Priority documents 0

Translations of priority documents 0

Statement of inventorship and right to grant of a patent (*Patents form 7/77*) 1 + 4 copies

Request for preliminary examination and search (*Patents Form 9/77*) 0

Request for Substantive Examination (*Patents Form 10/77*) 0

Any other documents (*please specify*) 0

11. I/We request the grant of a patent on the basis of this application

Signature



BERESFORD & Co

Date 3 August 2000

12. Name and daytime telephone number of person to contact in the United Kingdom

David SPROSTON

Tel: 020 7831 2290

Patents Form 7/77

Patents Act 1977

(Rule 15)

The  
Patent  
Office

**Statement of inventorship and of  
right to grant of a patent**

The Patent Office  
Cardiff Road  
Newport  
Gwent NP9 1RH

1. Your reference  
2636002/DS

2. Patent Application Number  
accompanying application reference 2636002

03 AUG 2000

3. Full name of the or each applicant  
Canon Kabushiki Kaisha

4. Title of the invention  
IMAGE PROCESSING APPARATUS

5. State how the applicant(s) derived the right from the inventor(s) to be granted a patent

By virtue of the employment of the inventors by Canon Research Europe Ltd, and by  
virtue of an agreement between Canon Research Europe Ltd and Canon Kabushiki  
Kaisha dated 1 January 1994.

6. How many, if any additional Patents Forms  
7/77 are attached to this form?

None.

7. I/We believe that the person(s) named over the page (and on any extra copies of this form) is/are the  
inventor(s) of the invention which the above patent application relates to.

Signature

  
BERESFORD & Co

Date 3 August 2000

8. Name and daytime telephone number of  
person to contact in the United Kingdom

David SPROSTON

Tel: 020 7831 2290

Patents Form 7/77

HASLAM; Jane  
7 Moreton Place  
Twynersh Avenue  
Chertsey  
Surrey  
KT16 9DE  
UNITED KINGDOM

9774730001

ROWE; Simon Michael  
c/o CANON RESEARCH CENTRE EUROPE LTD  
1 Occam Court, Occam Road  
Surrey Research Park  
Guildford  
Surrey GU2 5YJ  
UNITED KINGDOM

7736820001

TAYLOR; Richard Ian  
c/o CANON RESEARCH CENTRE EUROPE LTD  
1 Occam Court, Occam Road  
Surrey Research Park  
Guildford  
Surrey GU2 5YJ  
UNITED KINGDOM

7486707001

LYONS; Alexander Ralph  
c/o CANON RESEARCH CENTRE EUROPE LTD  
1 Occam Court, Occam Road  
Surrey Research Park  
Guildford  
Surrey GU2 5YJ  
UNITED KINGDOM

7486697001

## IMAGE PROCESSING APPARATUS

The present invention relates to an image processing apparatus and method.

5 It is known to create three dimensional computer models of real objects based on the input of image data in the form of a series of image frames which may be derived from a series of photographs taken from different camera positions or from a video recording taken from a moving camera. It is also known for such modelling techniques to require a user to identify coordinates in successive images of matching  
10 points, the input coordinates of matching points then being processed to create or refine the model, for example by calculating the positions in the coordinate system of the model from which the successive images were viewed by the camera and the three dimensional positions of the model points corresponding to the matched points.

15 This matching process of entering coordinates typically involves the user being presented on a display screen with a pair of successive images, for example in side by side relationship, and the user then being prompted to use a pointing device such as a computer mouse to move a cursor onto each selected image point and enter the coordinates of the point simply by actuating the pointing device, i.e. clicking the  
20 mouse, when the cursor is judged visually to be at the precise location of the image point selected.

It is also known to provide variable magnification of the displayed image as a whole in order to enable a user to zoom in on a portion of a displayed image of interest,  
25 thereby improving the accuracy with which the cursor position can be located prior to clicking the mouse.

It is also known to provide a portion of the display area with an enhanced magnification, typically referred to as a magnifying glass window, which can be moved under user actuation or selected by user actuation to provide localised enhanced magnification of the area of interest.

5

A problem exists in such known systems in that selection and control of the variable magnification facility requires additional actuation by the user of a keyboard or of the pointing device, thereby increasing complexity of operation and the amount of time required to complete the matching process.

10

Similar problems exist in processing image data for other purposes where it is required to repeatedly select a point within one frame and then select an associated point in a second frame with as much accuracy as possible in positioning the cursor in each case over the selected point.

15

A first aspect of the present invention seeks to provide an improved apparatus and method of processing such image data.

20

A further aspect of the present invention is concerned with the manner in which frames of the image data are selected when a user decides that it is necessary to update model data, either by adding further detail or correcting existing data, usually in respect of a particular localised feature of the model. If for example the model is to be updated by entering matching points between two frames of image data, the user must locate a pair of suitable image frames which present the relevant feature to the best advantage. Similarly, if data is to be corrected, the best view of the feature needs to be presented to the user in a frame of the image data for comparison with the

25



model image.

A further aspect of the present invention therefore seeks to provide an improved method and apparatus allowing the most appropriate camera images to be selected  
5 and displayed for use in the updating procedure.

According to the present invention there is disclosed a method of operating an apparatus for processing image data in accordance with user selected co-ordinates of displayed images representative of said image data; the apparatus performing the steps  
10 of;

displaying a first image representative of a first frame selected from said image data;

receiving pointing signals responsive to user actuation of a pointing device and displaying a cursor in the first image indicating an image point at a cursor position  
15 controlled by the pointing signals such that the cursor position is updated to track movement of the pointing device;

generating magnified image data representative of a first magnified image of a portion of the first image local to the cursor position and in fixed relationship thereto, and continuously updating the magnified image data in response to changes  
20 in the cursor position;

displaying the first magnified image simultaneously with the first image together with fiducial means indicating an image point in the first magnified image corresponding to the image point indicated in the first image at the cursor position;  
and

25 receiving a selection signal responsive to user actuation of said pointing device and representative of co-ordinates of a first selected point in the first image indicated

by the current cursor position.

Preferably the method further includes the step of displaying a second image representative of a second frame of said image data;

5           receiving pointing signals responsive to user actuation of the pointing device and displaying the cursor in the second image indicating an image point at a cursor position controlled by the pointing signals such that the cursor position is updated to track movement of the pointing device;

          generating magnified image data representative of a second magnified image  
10       of a portion of the second image local to the cursor position and in fixed relationship thereto, and continuously updating the magnified image data in response to changes in the cursor position;

          displaying the second magnified image simultaneously with the second image with second fiducial means indicating an image point in the second magnified image  
15       corresponding to the image point indicated in the second image at the cursor position; and

          receiving a selection signal responsive to user actuation of said pointing device and representative of co-ordinates of a second selected point in the second image indicated by the current cursor position.

20           According to a further aspect of the present invention there is disclosed a method of operating an apparatus for generating model data representative of a model in a three dimensional space of an object from input signals representative of a set of images of the object taken from a plurality of respective camera positions, the apparatus performing the steps of;

25           displaying a model image derived from the model data and comprising a plurality of primitives for viewing by a user;

receiving at least one primitive selection signal responsive to user actuation of an input means whereby each primitive selection signal identifies a respective selected primitive of the model;

5 defining a plurality of virtual cameras in the three dimensional space having positions and look directions relative to the model which correspond substantially to those of the respective actual cameras relative to the object;

evaluating which of the virtual cameras is an optimum virtual camera for generating a view of the selected primitives;

10 identifying from the camera images a first camera image of the plurality of camera images taken from a camera position corresponding to the optimum viewpoint.

In a preferred embodiment, the primitives are facets and the evaluating step calculates aspect measurements representative of the visibility of the facet when viewed in the look direction of each virtual camera. An alternative evaluating step calculates areas of the facet when viewed in projection in the look direction of each of the virtual cameras. In each case, the results of calculation are analysed to determine an optimum virtual camera and a complementary virtual camera so that a pair of camera images may be selected for display.

20

Preferred embodiments of the present invention will now be described by way of example only and with reference to the accompanying drawings of which;

Figure 1 is a schematic representation of a system for processing image data;

25

Figure 2 is a schematic representation of the apparatus of the present invention

including a processor having a display and pointing device for use in the system of Figure 1;

Figure 3 is a schematic representation of images displayed in the display screen of Figure 2 in accordance with the first aspect of the present invention, showing a first phase of operation in which a cursor is positioned in a first image;

Figure 4 is a further view of the display of Figure 3 showing a second phase in which the cursor is positioned in a second image;

10

Figure 5 is a schematic flowchart illustrating the first phase of operation;

Figure 6 is a schematic flowchart illustrating a second phase of operation;

15 Figure 7 is a schematic representation of a further phase of operation in which image points are matched in a third image; and

Figure 8 is a schematic representation, in a further aspect of the present invention, showing the initial orientation of a model image;

20

Figure 9 is a schematic representation of selection of a facet in the model image of Figure 8;

Figure 10 is a schematic representation of a display of the model image of Figures 8 and 9 in which multiple facets have been selected and camera images corresponding to an optimum view and a complementary view are displayed in conjunction with the

25

model image;

Figure 11 is a schematic diagram illustrating the relative position of virtual cameras relative to the model in three dimensions;

5

Figure 12 is a diagram illustrating the relationship between unit vectors used in an aspect measurement calculation;

10

Figure 13 is a diagram illustrating a projected area of a facet for use in visible area measurement;

Figure 14 is a graphical representation of aspect measurement for a given facet and for a plurality of virtual cameras;

15

Figure 15 is a graphical representation showing the frequency with which virtual cameras are selected as candidate virtual cameras for the selected set of facets;

20

Figure 16 is a schematic illustration of updating model data by the selection of matching points in camera images;

Figure 17A is a schematic illustration of updating model data using a drag and drop technique;

25

Figure 17B is a further illustration of the drag and drop technique, showing movement of a model point;

Figure 18A and 18B is a flowchart illustrating operation of the apparatus to select

camera images and update the image data;

Figure 19 is a flowchart illustrating selection of an optimum camera images;

Figure 20 is a flowchart illustrating determination of candidate virtual cameras;

Figure 21 is a flowchart illustrating the determination of the optimum virtual camera;

5 Figure 22 is a flowchart illustrating the determination of the optimum virtual camera based on viewable area measurements; and

Figure 23 is a flowchart illustrating an alternative method for updating model data using a drag and drop technique.

10 Figure 1 schematically shows the components of a modular system in which the present invention may be embodied.

These components can be effected as processor-implemented instructions, hardware or a combination thereof.

15

Referring to Figure 1, the components are arranged to process data defining images (still or moving) of one or more objects in order to generate data defining a three-dimensional computer model of the object(s).

20 The input image data may be received in a variety of ways, such as directly from one or more digital cameras, via a storage device such as a disk or CD ROM, by digitisation of photographs using a scanner, or by downloading image data from a database, for example via a datalink such as the Internet, etc.

25 The generated 3D model data may be used to: display an image of the object(s) from a desired viewing position; control manufacturing equipment to manufacture a model

of the object(s), for example by controlling cutting apparatus to cut material to the appropriate dimensions; perform processing to recognise the object(s), for example by comparing it to data stored in a database; carry out processing to measure the object(s), for example by taking absolute measurements to record the size of the object(s), or by comparing the model with models of the object(s) previously generated to determine changes therebetween; carry out processing so as to control a robot to navigate around the object(s); store information in a geographic information system (GIS) or other topographic database; or transmit the object data representing the model to a remote processing device for any such processing, either on a storage device or as a signal (for example, the data may be transmitted in virtual reality modelling language (VRML) format over the Internet, enabling it to be processed by a WWW browser); etc.

The feature detection and matching module 2 is arranged to receive image data recorded by a still camera from different positions relative to the object(s) (the different positions being achieved by moving the camera and/or the object(s)). The received data is then processed in order to match features within the different images (that is, to identify points in the images which correspond to the same physical point on the object(s)).

20

A further feature detection and tracking module 4 is arranged to receive image data recorded by a video camera as the relative positions of the camera and object(s) are changed (by moving the video camera and/or the object(s)). As in the feature detection and matching module 2, the feature detection and tracking module 4 detects features, such as corners, in the images. However, the feature detection and tracking module 4 then tracks the detected features between frames of image data in order to

determine the positions of the features in other images.

The camera position calculation module 6 is arranged to use the features matched across images by the feature detection and matching module 2 or the feature detection and tracking module 4 to calculate the transformation between the camera positions at which the images were recorded and hence determine the orientation and position of the camera focal plane when each image was recorded.

The feature detection and matching module 2 and the camera position calculation module 6 may be arranged to perform processing in an iterative manner. That is, using camera positions and orientations calculated by the camera position calculation module 6, the feature detection and matching module 2 may detect and match further features in the images using epipolar geometry in a conventional manner, and the further matched features may then be used by the camera position calculation module 6 to recalculate the camera positions and orientations.

If the positions at which the images were recorded are already known, then, as indicated by arrow 8 in Figure 1, the image data need not be processed by the feature detection and matching module 2, the feature detection and tracking module 4, or the camera position calculation module 6. For example, the images may be recorded by mounting a number of cameras on a calibrated rig arranged to hold the cameras in known positions relative to the object(s).

Alternatively, it is possible to determine the positions of a plurality of cameras relative to the object(s) by adding calibration markers to the object(s) and calculating the positions of the cameras from the positions of the calibration markers in images



recorded by the cameras. The calibration markers may comprise patterns of light projected onto the object(s). Camera calibration module 10 is therefore provided to receive image data from a plurality of cameras at fixed positions showing the object(s) together with calibration markers, and to process the data to determine the positions  
5 of the cameras. A preferred method of calculating the positions of the cameras (and also internal parameters of each camera, such as the focal length etc) is described in "Calibrating and 3D Modelling with a Multi-Camera System" by Wiles and Davison in 1999 IEEE Workshop on Multi-View Modelling and Analysis of Visual Scenes, ISBN 0769501109.

10

The 3D object surface generation module 12 is arranged to receive image data showing the object(s) and data defining the positions at which the images were recorded, and to process the data to generate a 3D computer model representing the actual surface(s) of the object(s), such as a polygon mesh model.

15

The texture data generation module 14 is arranged to generate texture data for rendering onto the surface model produced by the 3D object surface generation module 12. The texture data is generated from the input image data showing the object(s).

20

Techniques that can be used to perform the processing in the modules shown in Figure 1 are described in EP-A-0898245, EP-A-0901105, pending US applications 09/129077, 09/129079 and 09/129080, the full contents of which are incorporated herein by cross-reference, and also Annex A.

25

The present invention may be embodied in particular as part of the feature detection

and matching module 2 (although it has applicability in other applications, as will be described later).

Figure 2 illustrates a display monitor 20 having a display screen 21 on which are  
5 displayed first and second images 22 and 23. A processor 24 programmed with  
program code for creating a three dimensional computer model is connected to drive  
the display monitor 20 and receives pointing signals 25 from a computer mouse 26  
actuated by the user. The selection of frames of image data for providing the first and  
second images 22 and 23 may be made manually by the user or automatically by the  
10 processor 24 as described below with reference to Figures 8 to 23.

Additional data may also be input to the processor 24 via a keyboard 27. Software  
for operating the processor 24 is input to the processor from a portable storage  
medium in the form of a floppy disc 28 via a disc drive 29.

15

Figure 3 illustrates in greater detail the first and second images 22 and 23 displayed  
in the display screen 21, Figure 3 in particular showing a first phase of operation in  
which a cursor 30 is positioned within the first image. The cursor 30 is displayed by  
the display screen 21 at a position determined by movement of the mouse 26.

20

As shown in Figure 3, the first and second images 22 and 23 represent successive first  
and second frames of camera views of a real object, in this case a house, the camera  
views being from different camera positions.

25 The processor 24 causes the display monitor 20 to present the images of Figure 3 in  
response to user selection of a point matching mode, the interactive selection of

program operating modes by the user being by use of the computer mouse 26 and a menu of icons 48 displayed in a peripheral portion of the display screen 21.

During the first phase shown in Figure 3, the user selects visually an image point 31, in this example being an apex formed at the intersection of roof surfaces and an end wall of the house, and manipulates the mouse 26 to move the cursor 30 into a region of the first image proximate to the image point 31.

The first image 22 is displayed within a rectangular image window 33 which is partially overlaid by a first magnified image window 34. The first magnified image window 34 is square in shape and overlays the upper left hand corner of the image window 33. The first magnified image window 34 includes a graticule 35 in the form of horizontal and vertical cross wires intersecting at the centre of the first magnified image window.

A first magnified image 36 is displayed within the first magnified image window 34 and corresponds to a localised portion 32 of the first image 22, centred on the cursor position, and magnified to a sufficient magnitude to allow detail within the localised portion to be viewed more clearly by the user and to allow better resolution of any misalignment between the visually selected image point 31 and the image point corresponding to the current position of the cursor 30.

The processor 24 controls the display monitor 20 such that the first magnified image 36 is continuously displayed during a first phase of operation during which a point is to be selected in the first image. An enlarged view of the localised portion 32 is displayed, the image features displayed being determined instantaneously to be local

to the position of the cursor 30, it being apparent therefore that any movement of the cursor relative to the first image is accompanied by movement of image features within the first magnified image relative to the fixed graticule 35. The graticule 35 thereby serves as a fiducial means pointing to an image point 37 in the first magnified  
5 image corresponding to the same image feature as the image point 31 at the position of the cursor 30.

The first phase of operation ends when the user determines that the cursor 30 and graticule 35 are correctly aligned with the desired image point 37 for selection and the  
10 user actuates the pointing device, i.e. clicks the mouse 26, to generate a selection signal interpreted by the processor 24 as being representative of coordinates of a first selected point in the first image.

The processor thereafter freezes the first magnified image 36 within the first magnified  
15 image window 34 so that it continues to indicate alignment between the graticule 35 and the first selected point 37 irrespective of subsequent mouse movement. The processor 24 also generates an indicator 46 displayed in the first image 22 at the coordinates of the first selected point.

20 The user then operates the apparatus in a second phase illustrated in Figure 4 in which the cursor 30 is moved into the second image 23 with the intention of the user completing the matching process by selecting a second point corresponding to the same image feature as the first selected point 37 in the first image. The user visually identifies the feature of the apex in the house from the different view of the house  
25 shown in the second image and, as shown in Figure 4, moves the mouse 26 to position the cursor 30 in a region of the second image which is local to the apex.

The second image 23 is displayed within a second image window 41 which is rectangular in shape and which is overlaid at a top left hand corner by a second magnified image window 42 of similar square shape to the first magnified image window and similarly including a graticule 44 in the form of intersecting crosswires.

5

The display monitor 20 is controlled by the processor 24 to display within the second magnified image window 42, after commencement of the second phase, a second magnified image 43 corresponding to an enlargement of a localised portion 40 instantaneously determined to be local to the cursor 30 within the second image 23.

10

In this way, movement of the cursor 30 is accompanied by a change in view within the second magnified image window 42 so that the precise cursor position relative to the visually selected feature in the second image can be refined by viewing within the second magnified image window. Alignment is completed when the intersection of the cross wires of the graticule 44 is coincident with the selected feature and a second selected image point 45 is determined by actuating the pointing device, i.e. clicking the mouse.

15

The processor 24 interprets receiving a selection signal resulting from the mouse click as being representative of coordinates of the second selected image point indicated by the current cursor position, as confirmed by coincidence of the image feature with the graticule 44 in the second magnified image window 42.

20

The processor 24 thereafter controls the display monitor 20 to freeze the view displayed in the second magnified image window 42. Coordinates of the matching points defined by the first and second selected image points 37 and 45 are then

25

processed by the processor 24 to generate three dimensional model data for the model. In the system of Figure 1, this process is performed by the camera position calculation module 6 and the 3D surface generation model 12. Additional pairs of matching points may then be input in subsequent steps, each subsequent step  
5 comprising a respective first phase and second phase as described above.

To commence the matching for an additional pair of points, the user moves the cursor 30 back into the first image 22 to commence the first phase and the processor 24 then causes the first magnified image 36 to be unfrozen and to vary according to cursor  
10 position in the manner described above.

The method steps performed in the above process described with reference to Figures 3 and 4 are summarised in Figures 5 and 6 in which those steps performed by the user are shown separated from those steps performed by the apparatus by a broken line representing the interface 49 defined by the display screen 21 and user input devices  
15 including the mouse 26.

At step 50, the user selects the mode of operation which in this example is a matching mode for selecting matching points. The processor 24 receives the mode selection  
20 signal at step 51 displays at step 52 the first and second images 22 and 23 (as shown in Figure 3) and at step 53 the user views the images and decides upon a suitable image feature.

At step 54, the user actuates the pointing device, i.e. moves the mouse, to designate  
25 to a first approximation the position of the first image point 31 corresponding to the selected feature. At step 55, the processor receives the pointing signals resulting from

actuation of the pointing device, causing the display to indicate the cursor position accordingly at a user controlled position 30 within the first image.

At step 56, the processor causes the display to present a first magnified image 36 in  
5 the first magnified image window 34 so as to be continuously updated to be centred on the cursor coordinates.

At step 57, the user views the first magnified image 36 and refines the cursor position by viewing the magnified image. When finally the user is satisfied that the desired  
10 image feature is coincident with the intersecting crosswires of the graticule 35, the user actuates the selection switch of the computer mouse 26.

At step 58, the processor identifies the image coordinates at the selected position and freezes the view displayed in the first magnifier window.

15

The second phase illustrated schematically at Figure 6 then commences in which the user at step 60 actuates the mouse 26 to move the cursor into the second image 23 and, to a first approximation, aligns the cursor 30 with the matching image feature in the second image 23.

20

At step 61, the processor receives pointing signals corresponding to mouse movement and causes the display to display the cursor 30 at the user controlled position within the second image 23.

25 At step 62, a magnified view is displayed in the second magnified image window 42, a magnified image being displayed of the localised portion 40 of the second image

centred on the cursor coordinates.

At step 63, the user refines the pointer position using the second magnified image window 42 and actuates the selection switch of the mouse when the crosswires of the graticule 44 intersect precisely at the location of the matching image feature as viewed in the second magnified image 43.

At step 64, the processor identifies from selection signals generated by the mouse actuation the image coordinates of the selected matching position in the second image and fixes the magnified image displayed in the second magnified image window. At step 65, the processor stores the matched coordinates from the first and second images in a database of matched image points.

The next subsequent step of matching a pair of points then commences by returning to step 54 described above until the procedure is ultimately terminated by either the processor indicating that sufficient points have been matched or by the user selecting a different mode using a different one of the mode selecting icons 48.

By using the above apparatus and method, a user may rapidly enter successive pairs of matching points with the advantage of having a magnified view of the localised area of interest but with the minimum amount of actuation of the computer mouse since a single click of the mouse is required to select each one of the matching points. No further actuation of keyboard or mouse is needed to initiate generation of the required magnified view.

The matching procedure implemented by the feature detection and matching module



of the system of Figure 1 may in some circumstances require matching points to be identified in more than two images. A situation may then arise where the user wishes to identify in a third image a number of image points matched to a number of existing points for which matching co-ordinates have already been obtained in first and second  
5 images, using for example the method described above with reference to Figures 3, 4, 5 and 6.

In order to undertake the matching process to identify the points in the third image, the second and third images 71 and 72 are displayed side by side and the existing  
10 matched points are displayed in the second image by a series of indicators 70 in the form of crosses as illustrated in Figure 7. Magnified image windows 74 and 75 are provided in the image windows of the second and third images 71 and 72 respectively. The task of matching between the second and third images 71 and 72 shown in Figure 7 differs from the above described method with reference to Figures 3 and 4 since in  
15 the second image 71 the set of image points is predetermined by the previous matching step. To perform a matching process, the user selects one of the image points represented by the indicators 70 by placing the cursor on or adjacent to the image point and actuating the mouse. This pointing signal is detected by the processor 24 which then causes the displayed indicator 70 of the selected image point  
20 to be highlighted, for example by changing colour. In Figure 7, the selected point is highlighted by enclosing the indicator 70 by a circle 73. The magnified image window 74 then displays a magnified view of the second image local to the selected point.

The user then uses the mouse 26 to move the cursor 30 into the third image 72 and  
25 aligns the cursor 30 with the image feature corresponding to the selected point represented by the highlighted indicator 70,73 in the second image 71. Final

adjustment is made by viewing the magnified image within the magnified image window 75 in which the matching image point to be selected in the third image is identified by the location of the graticule 35 relative to the magnified image 75. The mouse 26 is then actuated by the user to provide a selection signal resulting in the input of co-ordinates to the model of matching image points in the second and third images 71 and 72. Matched points in the third image may be represented by indicators (not shown) as a guide to identifying which points in the second image remain to be matched.

Alternative embodiments are envisaged within the scope of the present invention including for example the use of alternative pointing devices such as a joystick or touch pad. Although in the preferred embodiment of Figures 2 to 7 the magnified image 74, 75 overlays a fixed portion of the displayed image, an alternative arrangement allows the operator to select the position of the magnified image window during an initial configuring step, the magnified image window thereafter remaining fixed in position. Such a configuring step may be advantageous where point matching is required in a peripheral portion of the image which might otherwise be hidden.

The graticule 35 within the magnified image window may alternatively be replaced by a stationary cursor, white spot or coloured spot, or any other fiducial means for identifying a fixed position within the magnified window.

The apparatus of the above embodiment may conveniently be constituted by a desktop computer operated by a computer program for operating the above described method steps in accordance with program code stored in the computer. The program code may be stored in a portable storage medium such as a CD ROM, floppy discs or

optical disc, represented generally by reference 28 in Figure 2.

An aspect of the present invention thus provides such a storage medium 28 storing processor implementable instructions for controlling a processor 24 to carry out the method described above.

Further, the computer program can be obtained in electronic form for example by downloading the code over a network such as the Internet. In Figure 2, a modem 38 suitable for such downloading is represented schematically.

Thus, in accordance with another aspect of present invention, there is provided an electrical signal 39 (Figure 2) carrying processor implementable instructions for controlling the processor 24 to carry out the method described above.

Further embodiments of the present invention are envisaged in which for example a series of points in a displayed image are selected by a user and co-ordinates of the selected points are input to a processor 24 with the aid of a magnified image as described above. Such alternatives include methods of categorising images such as fingerprint analysis and aerial photograph interpretation for use in cartography.

A further aspect of the present invention will now be illustrated by the following embodiments. This aspect of the invention may be used in the modular system of Figure 1 as described above and using the apparatus of Figure 2 including processor 24, display monitor 20 and computer mouse 26 actuated by the user.

As in the preceding embodiments, the processor 24 is programmed with program

code for creating a three-dimensional computer model, the processor being connected to drive the display monitor 20 and receive pointing signals 25 from the computer mouse 26.

5 Additional data may also be input to the processor 24 via keyboard 27. Software for operating the processor 24 is input to the processor from a portable storage medium in the form of a floppy disc 28 via a disc drive 29 or may be input in the form of a signal 39 via a modem 38.

10 Once model data has been created by processing image data of a number of frames of camera images, it is often the case that the user may judge that the model data requires refinement, for example to add further detail relating to a specific feature of the model or to correct model data in the case of the model image providing an incorrect representation of the object.

15

Procedures for adding and correcting model data typically require the display monitor to display both the model image and one or more camera images, in each case showing the relevant feature of the model and the object, to allow the user to interactively input model data and view the result when translated into an updated

20 model image. Since the model data may be derived from a large number of frames of image data, manual selection by the user of the most appropriate frames of image data may be time consuming and may provide less than optimum results. In accordance with the following embodiment, the processor 24 is therefore programmed to provide automatic selection of the most appropriate camera images for this purpose.

25

Control of the process relies upon the interface provided by the display screen 21 and

the input of pointing and selecting signals using computer mouse 26, steps in the method being illustrated in Figure 18 in which a left hand column contains steps conducted by the user and a right hand column contains steps executed by the apparatus in the form of the processor 24 connected to the display screen 21, the columns being separated by a broken line representing the interface. During the following description, reference will be made to the method steps shown in Figure 18 in relation to the images displayed on the display screen as shown in Figures 8 to 10.

The user at step 180 initially selects a model display mode from a menu of available modes of operation represented by mode selecting icons 48 and, in response to receiving the mode selecting input, the apparatus displays a view of the model in a model image window 81 as illustrated in Figure 8. In Figure 8, the representation of the model image 80 is illustrated as an irregular shape with a surface formed of a number of triangular facets. This representation is a simplified schematic representation, the actual model image typically being visually identifiable with a real object and comprising a much larger number of facets, the model image being rendered to include surface texture emulating the object.

The user actuates the mouse 26 to rotate the model image 80, left/right mouse movement effecting rotation of the model image in longitude as indicated by arrow 82 and forward/reverse movement of the mouse effecting rotation of the model image in latitude as indicated by arrow 83. A second mode of movement may be selected to vary the size of the model image. Throughout the above image movements, a virtual viewpoint for viewing the model is defined such that the model is always viewed in a direction directed to the centre of the co-ordinate system of the model data.

As shown in Figure 18, after selecting a viewpoint for the model image, such that the model image generated by the apparatus corresponds to a selected view showing a feature of particular interest to the user, the user selects at step 181 a facet selection mode. In this mode, movement of the mouse 26 effects movement of a cursor 30 relative to the model image 80 and, as shown in Figure 9, clicking the mouse 26 provides a facet selecting input signal in response to which a selected facet 90 at the location of the cursor 30 is highlighted in the model image, as illustrated by the cross hatched area in Figure 9.

- 10 The user is thereby able to select facets identifying a particular feature of interest in respect of which model data requires refinement or correction.

The user repeats facet selection until a set of selected facets is accumulated, as shown in Figure 10 in which the set of selected facets 100 are shaded.

- 15 As illustrated in Figure 10, the apparatus responds at step 183 by automatically selecting first and second camera images 101 and 102 which are displayed in a camera image window 103, based upon a determination of the optimum view of the model derived from the input of selected facets 100 described above.

- 20 The first camera image 101 includes a first view 104 of a feature constituted by a prominence of a particular shape protruding from the irregular surface of the object shown in the camera image, a second view 105 of the feature being provided in the second camera image 102. If the user is not satisfied that the correct camera images have been displayed, further facets may be added to the set 100 by selecting further facets shown in the model image window 81.
- 25

Once the user is satisfied that the displayed first and second camera images 101 and 102 are the most appropriate camera images, the user then selects at step 182 a model updating mode as shown in Figure 18. The apparatus continues to display the model and camera images and responds to further user input by following an interactive  
5 updating procedure based on the displayed images such that the model data is updated. The updated model data is used to update the displayed model image, giving the user the opportunity to continue the updating procedure to progressively refine the model as required.

10 According to a preferred embodiment using "aspect measurements" defined below, step 183 of selecting camera images as shown in Figure 18 is illustrated in further detail in the flowchart of Figure 19 and will now be described with additional reference to Figures 11 to 15. For each facet  $f$  of the selected facets 100 selected and highlighted during the facet selection mode of operation referred to above, a  
15 respective set of aspect measurements  $M(f,i)$ ,  $i = 1$  to  $n$  is calculated, each aspect measurement of the set being representative of the visibility of the facet when viewed from a virtual camera  $L(i)$ .

Figure 11 illustrates schematically the relationship between the three-dimensional  
20 model 110 and the virtual cameras  $L(i)$ ,  $i = 1$  to  $n$ . Each of the virtual cameras  $L(i)$  is represented by co-ordinates in the three dimensional space of the model to represent a camera position as calculated by the camera position calculation module 6 of Figure 1 and a look direction represented in Figure 11 by look direction vectors  $\underline{L}(i)$  which represent the direction normal to the image plane of the camera  $L(i)$ . The  
25 term "virtual camera" in the present context therefore refers to the calculated positions in model space corresponding to actual camera positions relative to the

object being modelled.

The method of calculating the aspect measurement  $M(f,i)$  is illustrated in Figure 12 which shows the relationship between a facet  $f$  and one of the virtual cameras  $L(i)$ .

5 The extent to which the facet  $f$  is visible with respect to virtual camera  $L(i)$  is dependent on the relationship between the look direction of the virtual camera, as defined by unit vector  $\underline{L}$ , and a unit vector  $\underline{f}$  defined to be a unit vector normal to the plane of the facet  $f$ . Defining  $\underline{L'}$  to be parallel to and in an opposite direction to the unit vector  $\underline{L}$ , the scalar product  $\underline{f} \cdot \underline{L'}$  has a magnitude which is representative of the  
10 extent to which the facet is visible. For example, a facet which has a normal unit vector  $\underline{f}$  parallel to the look direction  $\underline{L}$  will be fully visible and the scalar product will be unity whereas a facet oriented such that the look direction  $\underline{L}$  is parallel to the plane of the facet will have minimum visibility and the scalar product will be zero.

15 Figure 14 illustrates graphically for a given facet  $f$  the variation of aspect measurement with  $i$ , the identifier of the virtual cameras. In the example of Figure 14, a maximum value of aspect measurement is obtained for a virtual camera identified by  $i = I$  so that camera  $L(I)$  is identified as being a candidate for the optimum virtual camera.

20 The selection of optimised camera images as summarised in Figure 19 therefore includes at step 191 the step of determining a candidate virtual camera for each facet, the candidate virtual camera being in each case a respective virtual camera  $L(I)$  for which the aspect measurement  $M(f,i)$  has a maximum value. This determination is repeated for each of the facets as illustrated in the flowchart of Figure 20 where the  
25 results of aspect measurement are accumulated until all of the selected facets have been processed.



The accumulated results for a given set of facets are illustrated in Figure 15 in histogram form, showing the frequency with which each of the virtual cameras is selected to be a candidate virtual camera in step 191.

- 5 The virtual camera for which this frequency is a maximum is identified from the accumulated results as being the optimum virtual camera, illustrated in Figure 15 to correspond to the value  $i = X$ .

10 In Figure 19 therefore, step 192 of determining the optimum virtual camera consists of identifying the maximum frequency from the accumulated results of step 191, thereby identifying virtual camera X from the candidate virtual cameras and thereby allowing the first camera image to be identified at step 193 by identifying the image data yielding the position and look direction data for virtual camera X.

- 15 The first camera image 101 as illustrated in Figure 10 corresponds to this image data. To obtain the second camera image 102, a second virtual camera must be identified at step 194 of Figure 19. A complementary virtual camera is therefore selected from the accumulated results of aspect measurement according to a predetermined protocol in which, for a frequency distribution as shown in Figure 15, the complementary
- 20 virtual camera corresponds to  $i = X+1$ , being the virtual camera for which the next highest frequency is obtained in the direction of increasing  $i$ .

- The predetermined protocol for determining the complementary virtual camera may take account of frequency distributions in which there are twin peaks or where there
- 25 are several virtual cameras having the same maximum frequency by selecting the first maximum to occur in the direction of increasing  $i$  as being the optimum virtual camera

and the second maximum frequency to occur in the direction of increasing  $i$  as indicating the complementary virtual camera.

5 The image data selected for the second camera image 102 is identified as corresponding to the complementary virtual camera image and the first and second camera images are then displayed side by side as illustrated in Figure 10 in the camera image window 103.

10 As indicated in Figure 18B, the user then selects at step 184 the model updating mode which in the example of the present embodiment will be described in terms of updating the model data in response to the input of matching points in the first and second camera images. This method therefore utilises aspects of the method described above with reference to Figures 3 to 7.

15 During the updating procedure, the user successively enters image co-ordinates using the computer mouse 26 as a pointing device in conjunction with the cursor 30, matched points in the first and second camera images 101 and 102 being used by the apparatus to develop further model data and produce an updated model image 80 therefrom.

20

The user may then refine the appearance of the model image 80 to match more closely the camera images 101, 102. In particular, by matching points in the first and second camera images surrounding the feature seen in views 104 and 105 respectively of Figure 10, the model data relating to the region of the selection facets 100 may then  
25 be refined.

Figure 16 illustrates schematically the process of entering matching points 160 and 161 in the first and second camera images 101 and 102 respectively, the model image 80 being updated in real time accordingly as the model data is updated. A first point 160 is entered by clicking the mouse when the cursor 30 is positioned at a required feature in the first camera image and a second point 161 is then entered in the second camera image 102 at a position judged by the user to match the image feature identified by the first point 160. This matched pair of points is then processed by the apparatus. Further pairs of matched points are subsequently entered and the model image is incrementally updated accordingly.

As illustrated in Figure 18, the process ends when the updating of the model data is judged at step 185 to be complete by the user.

An alternative method of calculating the optimum virtual camera based on visible area measurement will now be described with reference to Figure 13, the method being based on a viewable area measurement. For each facet of the selected set of facets, a surface area  $A$  and a unit vector  $\underline{f}$  normal to the facet are defined. For a given virtual camera  $L(i)$  having a look direction defined by unit vector  $\underline{L}$ , the viewable area 130 when viewed from the virtual camera in projection in the look direction is proportional both to the scalar product of the unit vectors and to the area; a viewable area measurement  $V(i)$  is therefore defined to be

$$V(i) = A[\underline{f}\underline{L}]$$

where the square brackets indicate modulus. The viewable area measurement is calculated for each of the selected facets with respect to the virtual camera and summed to provide a total viewable area measurement  $S(i)$ .

The calculation of total viewable area measurement is repeated for each of the virtual cameras  $i$  and the optimum virtual camera determined as being the virtual camera for which  $S(i)$  is a maximum. The first camera image 101 may thereby be identified from this determination of the optimum virtual camera by determining the frame of image data associated with this virtual camera. The second camera image 102 may then be identified by determining a complementary virtual camera by determining the maximum total viewable area measurement of the remaining virtual cameras. As in the case of the aspect measurement process, ambiguities caused by a plurality of cameras having the same measurement are resolved by selecting virtual cameras in the order of increasing  $i$ .

The method steps for the calculation of the optimum virtual camera described above are illustrated in the flowchart of Figure 22.

An alternative method for updating the model data using a "drag and drop" technique will now be described with reference to Figures 17A and 17B and the method steps in the flowchart of Figure 23.

As indicated in Figure 23, the user selects at step 230 a model updating mode in response to which the apparatus displays (step 231) a model image 80 as shown in Figure 17A in a model image window 81, and at the same time displays first and second camera images 101 and 102 in a camera image window 103. The first and second camera images 101 and 102 may be selected by any of the above described methods. The user then selects (step 232) a facet 170 in the model image 80 using the cursor 30 and mouse, the apparatus responding to the generation of the facet selection signal by displaying (step 233) pointers 171, 172 and 173 in the model image

80 at corners of the facet 170 to represent model data points which can be edited. Corresponding pointers 174, 175 and 176 are mapped into each of the camera images 101 and 102 at locations determined in accordance with the camera position and look direction information associated with these frames of the image data.

5

As shown in Figure 17A, the camera images 101 and 102 include a prominent feature 177, the apex of which is represented in the model image by pointer 172 which, as illustrated schematically in Figure 17A, is incorrectly positioned when compared with the camera images. The user then uses the mouse 26 and cursor 30 to manipulate (step 234) the position of the pointer 172 in the model image 80 using a "drag and drop" technique in which the mouse is actuated to select the pointer 172 and the mouse actuating key depressed while moving the mouse and cursor position to a revised position. The apparatus tracks this movement (step 235) and, on releasing the mouse, the pointer 172 then remains in its edited position. The user may decide (step 15 236) to carry out further editing, repeating steps 234 and 235 accordingly. The model data is updated in accordance with the edited positions. Although the movement of the pointer 172 defines movement of the model point in only two dimensions, the edited model point position can be determined by constraining movement to lie in a plane orthogonal to the direction in which the projection of the model is viewed to arrive at the model image 80. 20

The editing process is illustrated in Figure 17B in which the new position of the pointer 172 is shown in the model image. Throughout this editing process, the position of the corresponding pointers 175 in the camera images 101 and 102 are updated in real time so that the user may observe this movement until, as shown in 25 Figure 17B, these pointers are coincident with the apex of the feature 177. The model

data is thereby edited such that the model image represents more closely the prominent feature 177.

As illustrated in the flowchart of Figure 23, this editing procedure may be repeated  
5 by dragging and dropping further pointers from the same facet 170 or by selecting further facets to access additional pointers.

The above mentioned methods for selecting the optimum virtual camera in order to select the best camera image ensure that the above drag and drop editing process is  
10 carried out in the simplest and most effective manner since the best camera images are provided to the user for the editing procedure.

The apparatus of the above embodiment may conveniently be constituted by a desktop computer operated by a computer program for operating the above described method  
15 steps in accordance with program code stored in the computer. The program code may be stored in a portable storage medium such as a CD ROM, floppy discs or optical disc, represented generally by reference 28 in Figure 2.

An aspect of the present invention thus provides such a storage medium storing  
20 processor implementable instructions for controlling a processor to carry out the method described above.

Further, the computer program can be obtained in electronic form for example by downloading the code over a network such as the Internet. In Figure 2, a modem  
25 suitable for such downloading is represented schematically.

Thus, in accordance with another aspect of present invention, there is provided an electrical signal 39 (Figure 2) carrying processor implementable instructions for controlling the processor to carry out the method described above.

- 5 Further embodiments of the present invention are envisaged in which for example the display of the model image may be other than a rendered image and may for example be in the form of a wire frame.

- 10 The embodiments described with reference to Figures 8 to 23 refer to the selection of facets in the model image. More generally, the invention is applicable to the selection of any appropriate primitives in the model, such as for example, polygonal facets of more than three sides, lines or three-dimensional elements, and corresponding methods using such primitives are intended to fall within the scope of the present invention by appropriate modification to the above described  
15 embodiments.

- Similarly, in the drag and drop method described above with reference to Figures 17A and 17B, other primitives may be moved by the drag and drop technique, for example the entire facet may be moved in a manner which retains its shape or a line may be  
20 translated from one position to another. The drag and drop technique may also incorporate rotational movement for those primitives in respect of which such rotation would be appropriate.

- In the above described technique of matching points as shown in Figure 16, a  
25 magnified image window of the type illustrated in Figure 3 may additionally be provided in each of the camera images in order to assist the operator in accurate

cursor movement, using the method described above with reference to Figures 3 to 7.

The present application incorporates by cross-reference the full contents of the applicant's following co-pending applications: 9927876.4 (attorney ref: 2635901); 9927875.6 (attorney ref: 2636501); 9927906.9 (attorney ref: 2641601); 9927907.7 (attorney ref: 2641701); 9927909.3 (attorney ref: 2641901); 9929957.0 (attorney ref: 2641001); 0001479.5 (attorney ref: 2636201) and 0018492.9 (attorney ref: 2635801).



## ANNEX A

### 1. CORNER DETECTION

#### 5 1.1 Summary

This process described below calculates corner points, to sub-pixel accuracy, from a single grey scale or colour image. It does this by first detecting edge boundaries in the image and then choosing corner points to be points where a strong edge changes  
10 direction rapidly. The method is based on the facet model of corner detection, described in Haralick and Shapiro<sup>1</sup>.

#### 1.2 Algorithm

15 The algorithm has four stages:

- (1) Create grey scale image (if necessary);
- (2) Calculate edge strengths and directions;
- (3) Calculate edge boundaries;
- 20 (4) Calculate corner points.

##### 1.2.1 Create grey scale image

The corner detection method works on grey scale images. For colour images, the colour values are first converted to floating point grey scale values using the formula:

$$grey\_scale = (0.3 \times red) + (0.59 \times green) + (0.11 \times blue)$$

5

....A-1

This is the standard definition of brightness as defined by NTSC and described in Foley and van Dam<sup>ii</sup>.

10

#### 1.2.2 Calculate edge strengths and directions

The edge strengths and directions are calculated using the 7x7 integrated directional derivative gradient operator discussed in section 8.9 of Haralick and Shapiro<sup>i</sup>.

15

The row and column forms of the derivative operator are both applied to each pixel in the grey scale image. The results are combined in the standard way to calculate the edge strength and edge direction at each pixel.

20 The output of this part of the algorithm is a complete derivative image.

#### 1.2.3 Calculate edge boundaries

25 The edge boundaries are calculated by using a zero crossing edge detection method based on a set of 5x5 kernels describing a bivariate cubic fit to the neighbourhood of

each pixel.

The edge boundary detection method places an edge at all pixels which are close to a negatively sloped zero crossing of the second directional derivative taken in the direction of the gradient, where the derivatives are defined using the bivariate cubic fit to the grey level surface. The subpixel location of the zero crossing is also stored along with the pixel location.

The method of edge boundary detection is described in more detail in section 8.8.4 of Haralick and Shapiro<sup>1</sup>.

#### 1.2.4 Calculate corner points

The corner points are calculated using a method which uses the edge boundaries calculated in the previous step.

Corners are associated with two conditions:

- (1) the occurrence of an edge boundary; and
- (2) significant changes in edge direction.

Each of the pixels on the edge boundary is tested for "cornerness" by considering two points equidistant to it along the tangent direction. If the change in the edge direction is greater than a given threshold then the point is labelled as a corner. This step is described in section 8.10.1 of Haralick and Shapiro<sup>1</sup>.

5

Finally the corners are sorted on the product of the edge strength magnitude and the change of edge direction. The top 200 corners which are separated by at least 5 pixels are output.

## 10 2. FEATURE TRACKING

### 2.1 Summary

This process described below tracks feature points (typically corners) across a  
15 sequence of grey scale or colour images.

The tracking method uses a constant image velocity Kalman filter to predict the motion of the corners, and a correlation based matcher to make the measurements of corner correspondences.

20

The method assumes that the motion of corners is smooth enough across the sequence of input images that a constant velocity Kalman filter is useful, and that corner

measurements and motion can be modelled by gaussians.

## 2.2 Algorithm

- 5     1)     Input corners from an image.
- 2)     Predict forward using Kalman filter.
- 3)     If the position uncertainty of the predicted corner is greater than a threshold,  
10          $\Delta$ , as measured by the state positional variance, drop the corner from the list  
             of currently tracked corners.
- 4)     Input a new image from the sequence.
- 15     5)     For each of the currently tracked corners:
  - a)     search a window in the new image for pixels which match the corner;
  - b)     update the corresponding Kalman filter, using any new observations  
         (i.e. matches).
- 20     6)     Input the corners from the new image as new points to be tracked (first,  
             filtering them to remove any which are too close to existing tracked points).



7) Go back to (2)

### 2.2.1 Prediction

- 5 This uses the following standard Kalman filter equations for prediction, assuming a constant velocity and random uniform gaussian acceleration model for the dynamics:

$$X_{n+1} = \Theta_{n+1,n} X_n \quad \dots A-2$$

$$K_{n+1} = \Theta_{n+1,n} K_n \Theta_{n+1,n}^T + Q_n \quad \dots A-3$$

- where  $x$  is the 4D state of the system, (defined by the position and velocity vector of the corner),  $K$  is the state covariance matrix,  $\Theta$  is the transition matrix, and  $Q$  is the process covariance matrix.
- 10

In this model, the transition matrix and process covariance matrix are constant and have the following values:

$$\Theta_{n+1,n} = \begin{pmatrix} I & I \\ 0 & I \end{pmatrix} \quad \dots A-4$$

$$Q_n = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_{vI}^2 \end{pmatrix}$$

....A-5

5

### 2.2.2 Searching and matching

This uses the positional uncertainty (given by the top two diagonal elements of the state covariance matrix, K) to define a region in which to search for new measurements (i.e. a range gate).

10

The range gate is a rectangular region of dimensions:

$$\Delta x = \sqrt{K_{11}}, \Delta y = \sqrt{K_{22}}$$

....A-6

15

The correlation score between a window around the previously measured corner and each of the pixels in the range gate is calculated.

20 The two top correlation scores are kept.

If the top correlation score is larger than a threshold,  $C_0$ , and the difference between the two top correlation scores is larger than a threshold  $\Delta C$ , then the pixel with the top correlation score is kept as the latest measurement.



### 2.2.3 Update

The measurement is used to update the Kalman filter in the standard way:

$$G = KH^T(HKH^T + R)^{-1} \quad \dots A-7$$

5

$$x \leftarrow x + G(\hat{x} - Hx) \quad \dots A-8$$

$$K \leftarrow (I - GH)K \quad \dots A-9$$

where G is the Kalman gain, H is the measurement matrix,

10 and R is the measurement covariance matrix.

In this implementation, the measurement matrix and measurement covariance matrix are both constant, being given by:

$$H = (I \ 0) \quad \dots A-10$$

15

$$R = \sigma^2 I \quad \dots A-11$$

### 2.2.4 Parameters

20 The parameters of the algorithm are:

Initial conditions:  $x_0$  and  $K_0$ .

Process velocity variance:  $\sigma_v^2$ .

Measurement variance:  $\sigma^2$ .

Position uncertainty threshold for loss of track:  $\Delta$ .

Covariance threshold:  $C_0$ .

Matching ambiguity threshold:  $\Delta C$ .

- 5 For the initial conditions, the position of the first corner measurement and zero velocity are used, with an initial covariance matrix of the form:

$$K_0 = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_0^2 I \end{pmatrix} \quad \dots A-12$$

- 10  $\sigma_0^2$  is set to  $\sigma_0^2 = 200(\text{pixels/frame})^2$ .

The algorithm's behaviour over a long sequence is anyway not too dependent on the initial conditions.

- 15 The process velocity variance is set to the fixed value of 50 (pixels/frame)<sup>2</sup>. The process velocity variance would have to be increased above this for a hand-held sequence. In fact it is straightforward to obtain a reasonable value for the process velocity variance adaptively.

- 20 The measurement variance is obtained from the following model:

$$\sigma^2 = (rK + a) \quad \dots A-13$$

where  $K = \sqrt{(K_{11}K_{22})}$  is a measure of the positional uncertainty,  $r$  is a parameter related to the likelihood of obtaining an outlier, and  $a$  is a parameter related to the measurement uncertainty of inliers. " $r$ " and " $a$ " are set to  $r=0.1$  and  $a=1.0$ .

- 5 This model takes into account, in a heuristic way, the fact that it is more likely that an outlier will be obtained if the range gate is large.

The measurement variance (in fact the full measurement covariance matrix  $R$ ) could also be obtained from the behaviour of the auto-correlation in the neighbourhood of  
10 the measurement. However this would not take into account the likelihood of obtaining an outlier.

The remaining parameters are set to the values:  $\Delta=400$  pixels<sup>2</sup>,  $C_0=0.9$  and  $\Delta C=0.001$ .

### 15 3. 3D SURFACE GENERATION

#### 3.1 Architecture

In the method described below, it is assumed that the object can be segmented from  
20 the background in a set of images completely surrounding the object. Although this restricts the generality of the method, this constraint can often be arranged in practice, particularly for small objects.

The method consists of five processes, which are run consecutively:

- First, for all the images in which the camera positions and orientations have been calculated, the object is segmented from the background, using colour information. This produces a set of binary images, where the pixels are marked as being either object or background.
- The segmentations are used, together with the camera positions and orientations, to generate a voxel carving, consisting of a 3D grid of voxels enclosing the object. Each of the voxels is marked as being either object or empty space.
- The voxel carving is turned into a 3D surface triangulation, using a standard triangulation algorithm (marching cubes).
- The number of triangles is reduced substantially by passing the triangulation through a decimation process.
- Finally the triangulation is textured, using appropriate parts of the original images to provide the texturing on the triangles.

### 3.2 Segmentation

The aim of this process is to segment an object (in front of a reasonably homogeneous coloured background) in an image using colour information. The resulting binary  
5 image is used in voxel carving.

Two alternative methods are used:

Method 1: input a single RGB colour value representing the background  
10 colour - each RGB pixel in the image is examined and if the Euclidean distance to the background colour (in RGB space) is less than a specified threshold the pixel is labelled as background (BLACK).

Method 2: input a "blue" image containing a representative region of the  
15 background.

The algorithm has two stages:

- (1) Build a hash table of quantised background colours  
20
- (2) Use the table to segment each image.

Step 1) Build hash table

Go through each RGB pixel,  $p$ , in the "blue" background image.

- 5 Set  $q$  to be a quantised version of  $p$ . Explicitly:

$$q = (p + t/2)/t \quad \text{....A-14}$$

where  $t$  is a threshold determining how near RGB values need to be to background colours to be labelled as background.

- 10 The quantisation step has two effects:

- 1) reducing the number of RGB pixel values, thus increasing the efficiency of hashing;
- 15 2) defining the threshold for how close a RGB pixel has to be to a background colour pixel to be labelled as background.

$q$  is now added to a hash table (if not already in the table) using the (integer) hashing function

20

$$h(q) = (q_{red} \& 7) * 2^6 + (q_{green} \& 7) * 2^3 + (q_{blue} \& 7)$$

....A-15

That is, the 3 least significant bits of each colour field are used. This function is chosen to try and spread out the data into the available bins. Ideally each bin in the hash table has a small number of colour entries. Each quantised colour RGB triple is only added once to the table (the frequency of a value is irrelevant).

5

Step 2) Segment each image

Go through each RGB pixel,  $v$ , in each image.

10 Set  $w$  to be the quantised version of  $v$  as before.

To decide whether  $w$  is in the hash table, explicitly look at all the entries in the bin with index  $h(w)$  and see if any of them are the same as  $w$ . If yes, then  $v$  is a background pixel - set the corresponding pixel in the output image to BLACK. If no  
15 then  $v$  is a foreground pixel - set the corresponding pixel in the output image to WHITE

Post Processing: For both methods a post process is performed to fill small holes and remove small isolated regions.

20

A median filter is used with a circular window. (A circular window is chosen to avoid biasing the result in the  $x$  or  $y$  directions).

Build a circular mask of radius  $r$ . Explicitly store the start and end values for each scan line on the circle.

Go through each pixel in the binary image.

5

Place the centre of the mask on the current pixel. Count the number of BLACK pixels and the number of WHITE pixels in the circular region.

If ( $\# \text{WHITE pixels} \geq \# \text{BLACK pixels}$ ) then set corresponding output pixel to  
10 WHITE. Otherwise output pixel is BLACK.

### 3.3 Voxel carving

The aim of this process is to produce a 3D voxel grid, enclosing the object, with each  
15 of the voxels marked as either object or empty space.

The input to the algorithm is:

- a set of binary segmentation images, each of which is associated with a camera  
20 position and orientation;

- 2 sets of 3D co-ordinates, ( $x_{\min}$ ,  $y_{\min}$ ,  $z_{\min}$ ) and ( $x_{\max}$ ,  $y_{\max}$ ,  $z_{\max}$ ),



describing the opposite vertices of a cube surrounding the object;

- a parameter,  $n$ , giving the number of voxels required in the voxel grid.

5 A pre-processing step calculates a suitable size for the voxels (they are cubes) and the 3D locations of the voxels, using  $n$ ,  $(x_{\min}, y_{\min}, z_{\min})$  and  $(x_{\max}, y_{\max}, z_{\max})$ .

Then, for each of the voxels in the grid, the mid-point of the voxel cube is projected into each of the segmentation images. If the projected point falls onto a pixel which  
10 is marked as background, on any of the images, then the corresponding voxel is marked as empty space, otherwise it is marked as belonging to the object.

Voxel carving is described further in "Rapid Octree Construction from Image Sequences" by R. Szeliski in CVGIP: Image Understanding, Volume 58, Number 1,  
15 July 1993, pages 23-32.

### 3.4 Marching cubes

The aim of the process is to produce a surface triangulation from a set of samples of  
20 an implicit function representing the surface (for instance a signed distance function).

In the case where the implicit function has been obtained from a voxel carve, the implicit function takes the value -1 for samples which are inside the object and +1 for

samples which are outside the object.

Marching cubes is an algorithm that takes a set of samples of an implicit surface (e.g. a signed distance function) sampled at regular intervals on a voxel grid, and extracts a triangulated surface mesh. Lorensen and Cline<sup>iii</sup> and Bloomenthal<sup>iv</sup> give details on the algorithm and its implementation.

The marching-cubes algorithm constructs a surface mesh by "marching" around the cubes while following the zero crossings of the implicit surface  $f(x)=0$ , adding to the triangulation as it goes. The signed distance allows the marching-cubes algorithm to interpolate the location of the surface with higher accuracy than the resolution of the volume grid. The marching cubes algorithm can be used as a continuation method (i.e. it finds an initial surface point and extends the surface from this point).

### 3.5 Decimation

15

The aim of the process is to reduce the number of triangles in the model, making the model more compact and therefore easier to load and render in real time.

The process reads in a triangular mesh and then randomly removes each vertex to see if the vertex contributes to the shape of the surface or not. (i.e. if the hole is filled, is the vertex a "long" way from the filled hole). Vertices which do not contribute to the shape are kept out of the triangulation. This results in fewer vertices (and hence

20

triangles) in the final model.

The algorithm is described below in pseudo-code.

```
5      INPUT
      Read in vertices
      Read in triples of vertex IDs making up triangles

      PROCESSING
10     Repeat NVERTEX times
          Choose a random vertex,  $V$ , which hasn't been chosen before
          Locate set of all triangles having  $V$  as a vertex,  $S$ 
          Order  $S$  so adjacent triangles are next to each other
          Re-triangulate triangle set, ignoring  $V$  (i.e. remove selected triangles
15          &  $V$  and then fill in hole)
          Find the maximum distance between  $V$  and the plane of each triangle
          If (distance < threshold)
              Discard  $V$  and keep new triangulation
          Else
20              Keep  $V$  and return to old triangulation

      OUTPUT
      Output list of kept vertices
      Output updated list of triangles
```

25

The process therefore combines adjacent triangles in the model produced by the marching cubes algorithm, if this can be done without introducing large errors into the model.

The selection of the vertices is carried out in a random order in order to avoid the effect of gradually eroding a large part of the surface by consecutively removing neighbouring vertices.

### 5    3.6    Further Surface Generation Techniques

Further techniques which may be employed to generate a 3D computer model of an object surface include voxel colouring, for example as described in "Photorealistic Scene Reconstruction by Voxel Coloring" by Seitz and Dyer in Proc. Conf. Computer Vision and Pattern Recognition 1997, p1067-1073, "Plenoptic Image Editing" by Seitz and Kutulakos in Proc. 6th International Conference on Computer Vision, pp 17-24, "What Do N Photographs Tell Us About 3D Shape?" by Kutulakos and Seitz in University of Rochester Computer Sciences Technical Report 680, January 1998, and "A Theory of Shape by Space Carving" by Kutulakos and Seitz in University of Rochester Computer Sciences Technical Report 692, May 1998.

## 4.    TEXTURING

The aim of the process is to texture each surface polygon (typically a triangle) with the most appropriate image texture. The output of the process is a VRML model of the surface, complete with texture co-ordinates.

The triangle having the largest projected area is a good triangle to use for texturing, as it is the triangle for which the texture will appear at highest resolution.

5 A good approximation to the triangle with the largest projected area, under the assumption that there is no substantial difference in scale between the different images, can be obtained in the following way.

10 For each surface triangle, the image "i" is found such that the triangle is the most front facing (i.e. having the greatest value for  $\hat{n}_i \cdot \hat{v}_i$ , where  $\hat{n}_i$  is the triangle normal and  $\hat{v}_i$  is the viewing direction for the "i" th camera). The vertices of the projected triangle are then used as texture co-ordinates in the resulting VRML model.

15 This technique can fail where there is a substantial amount of self-occlusion, or several objects occluding each other. This is because the technique does not take into account the fact that the object may occlude the selected triangle. However, in practice this does not appear to be much of a problem.

20 It has been found that, if every image is used for texturing then this can result in very large VRML models being produced. These can be cumbersome to load and render in real time. Therefore, in practice, a subset of images is used to texture the model. This subset may be specified in a configuration file.

References

- i R M Haralick and L G Shapiro: "Computer and Robot Vision Volume 1",  
Addison-Wesley, 1992, ISBN 0-201-10877-1 (v.1), section 8.
- 5
- ii J Foley, A van Dam, S Feiner and J Hughes: "Computer Graphics: Principles  
and Practice", Addison-Wesley, ISBN 0-201-12110-7.
- iii W.E. Lorensen and H.E.Cline: "Marching Cubes: A High Resolution 3D  
10 Surface Construction Algorithm", in Computer Graphics, SIGGRAPH 87  
proceedings, 21: 163-169, July 1987.
- iv J. Bloomenthal: "An Implicit Surface Polygonizer", Graphics Gems IV, AP  
Professional, 1994, ISBN 0123361559, pp 324-350.

CLAIMS

1. A method of operating an apparatus for processing image data in accordance with user selected co-ordinates of displayed images representative of said image data; the apparatus performing the steps of;

5 displaying a first image representative of a first frame selected from said image data;

receiving pointing signals responsive to user actuation of a pointing device and displaying a cursor in the first image indicating an image point at a cursor position controlled by the pointing signals such that the cursor position is updated to track  
10 movement of the pointing device;

generating magnified image data representative of a first magnified image of a portion of the first image local to the cursor position and in fixed relationship thereto, and continuously updating the magnified image data in response to changes in the cursor position;

15 displaying the first magnified image simultaneously with the first image together with fiducial means indicating an image point in the first magnified image corresponding to the image point indicated in the first image at the cursor position; and

receiving a selection signal responsive to user actuation of said pointing device  
20 and representative of co-ordinates of a first selected point in the first image indicated by the current cursor position.

2. A method as claimed in claim 1 wherein the step of displaying the first magnified image comprises displaying the first magnified image in a first window which overlays a fixed portion of the first image.

5 3. A method as claimed in any preceding claim wherein the step of displaying of the fiducial means comprises displaying a graticule.

4. A method as claimed in any preceding claim including the step of sampling the magnified image data at the time of receiving the selection signal, storing the sampled  
10 data and continuing to display the first magnified image as a static image corresponding to the stored image data.

5. A method as claimed in any preceding claim including the step of displaying a second image representative of a second frame of said image data;  
15 receiving pointing signals responsive to user actuation of the pointing device and displaying the cursor in the second image indicating an image point at a cursor position controlled by the pointing signals such that the cursor position is updated to track movement of the pointing device;

generating magnified image data representative of a second magnified image  
20 of a portion of the second image local to the cursor position and in fixed relationship thereto, and continuously updating the magnified image data in response to changes in the cursor position;



displaying the second magnified image simultaneously with the second image with second fiducial means indicating an image point in the second magnified image corresponding to the image point indicated in the second image at the cursor position; and

5           receiving a selection signal responsive to user actuation of said pointing device and representative of co-ordinates of a second selected point in the second image indicated by the current cursor position.

6.       A method as claimed in claim 5 wherein the second magnified image is  
10       displayed in a second window which overlays a fixed portion of the second image.

7.       A method as claimed in claim 5 including the step of storing co-ordinates of the first and second selected points constituting matching points in the first and second images respectively.

15       8.       A method as claimed in claim 7 including the step of processing the co-ordinates of the matching points to generate model data for a model in a three-dimensional space of an object represented in camera images from which said image data is derived.

20       9.       Apparatus for processing image data in accordance with user selected co-ordinates of displayed images representative of said image data; the apparatus

comprising;

display means operable to display a first image representative of a first frame selected from said image data;

5 pointing signal receiving means for receiving pointing signals responsive to user actuation of a pointing device and causing the display means to display a cursor in the first image indicating an image point at a cursor position controlled by the pointing signals such that the cursor position is updated to track movement of the pointing device;

10 generating means for generating magnified image data representative of a first magnified image of a portion of the first image local to the cursor position and in fixed relationship thereto, and for continuously updating the magnified image data in response to changes in the cursor position;

the display means being further operable to display the first magnified image simultaneously with the first image together with fiducial means indicating an image  
15 point in the first magnified image corresponding to the image point indicated in the first image at the cursor position; and

selection signal receiving means for receiving a selection signal responsive to user actuation of said pointing device in use and representative of co-ordinates of a first selected point in the first image indicated by the current cursor position.

20

10. Apparatus as claimed in claim 9 wherein the display means is operable to display the first magnified image in a first window which overlays a fixed portion of

the first image.

11. Apparatus as claimed in of claims 9 and 10 wherein the fiducial means comprises a graticule.

5

12. Apparatus as claimed in any of claims 9 to 11 including means for sampling the magnified image data at the time of receiving the selection signal, storing the sampled data and continuing to display the first magnified image as a static image corresponding to the stored image data.

10

13. Apparatus as claimed in any of claims 9 to 12 wherein the display means is operable to display a second image representative of a second frame of said image data;

the pointing signal receiving means being operable to receive pointing signals responsive to further user actuation of the pointing device and causing the display means to display the cursor in the second image indicating an image point at a cursor position controlled by the pointing signals such that the cursor position is updated to track movement of the pointing device;

the generating means being further operable to generate magnified image data representative of a second magnified image of a portion of the second image local to the cursor position and in fixed relationship thereto, and to continuously update the magnified image data in response to changes in the cursor position;

20

the display means being operable to display the second magnified image simultaneously with the second image with second fiducial means indicating an image point in the second magnified image corresponding to the image point indicated in the second image at the cursor position; and

5 the selection signal receiving means being operable to receive a selection signal responsive to user actuation of said pointing device and representative of co-ordinates of a second selected point in the second image indicated by the current cursor position.

10 14. Apparatus as claimed in claim 13 wherein the second magnified image is displayed in a second window which overlays a fixed portion of the second image.

15 15. Apparatus as claimed in claim 13 including means for storing co-ordinates of the first and second selected points constituting matching points in the first and second images respectively.

16. Apparatus as claimed in claim 15 including means for processing the co-ordinates of the matching points to generate model data for a model in a three dimensional space of an object represented in camera images from which said image data is derived.

20

17. A storage medium storing processor implementable instructions for controlling

a processor to carry out a method as claimed in any of claims 1 to 8.

18. An electrical signal carrying processor implementable instructions for controlling a processor to carry out a method as claimed in any of claims 1 to 8.

5

19. A computer program comprising processor implementable instructions for controlling a processor to carry out a method as claimed in any of claims 1 to 8.

20. A method of operating an apparatus for generating model data representative of a model in a three dimensional space of an object from input signals representative of a set of images of the object taken from a plurality of respective camera positions, the apparatus performing the steps of;

displaying a model image derived from the model data and comprising a plurality of primitives for viewing by a user;

15 receiving at least one primitive selection signal responsive to user actuation of an input means whereby each primitive selection signal identifies a respective selected primitive of the model;

defining a plurality of virtual cameras in the three dimensional space having positions and look directions relative to the model which correspond substantially to those of the respective actual cameras relative to the object;

20 evaluating which of the virtual cameras is an optimum virtual camera for generating a view of the selected primitives;

identifying from the camera images a first camera image of the plurality of camera images taken from a camera position corresponding to that of the optimum virtual camera.

- 5      21.      A method as claimed in claim 20 including the step of determining from the camera images a second camera image as being suitable for matching features in the first camera image and displaying the second camera image for comparison by the user with the first camera image.
- 10      22.      A method as claimed in claim 21 wherein the second camera image is taken from a camera position proximate to the optimum camera position.
- 15      23.      A method as claimed in any of claims 21 and 22 including the step of receiving feature matching selection signals representative of user matched points in the first and second camera images.
- 20      24.      A method as claimed in claim 23 including the step of generating updated model data to include additional detail corresponding to the received feature matching selection signals rendering the updated model data to generate an updated model image and displaying the updated model image.
25.      A method as claimed in any of claims 20 to 24 wherein the evaluating step

comprises;

calculating for a selected primitive an aspect measurement representative of the visibility of the primitive when viewed in projection in the look direction of one of the virtual cameras;

5           repeating the calculating step to obtain a respective aspect measurement for each of the virtual cameras;

comparing the aspect measurements for the selected primitive and determining a candidate virtual camera to be the virtual camera for which the corresponding aspect measurement is a maximum;

10           repeating the calculating, comparing and determining steps for each of the selected primitive whereby candidate virtual cameras are determined for each selected primitive; and

choosing the optimum virtual camera on the basis of the frequency with which virtual cameras are determined to be candidate virtual cameras.

15

26.    A method as claimed in claim 25 wherein the primitives comprise facets.

27.    A method as claimed in claim 26 wherein the calculation of the aspect measurement comprises, for a given facet and a given virtual camera, calculating a  
20    scalar product of a unit vector normal to the facet and a unit vector parallel to the look direction of the virtual camera.

28. A method as claimed in claim 26 wherein the calculation of aspect measurement comprises calculating, for a given facet and for a given virtual camera, an area of the facet when viewed in projection in the look direction of the virtual camera.

5

29. A method as claimed in any of claims 20 to 28 wherein the input means is a pointing means co-operable with a display means to provide input signals in the form of image co-ordinates of the displayed image.

10 30. A method as claimed in any of claims 20 to 29 including generating the displayed model image by rendering the image data.

31. Apparatus for generating model data representative of a model in a three dimensional space of an object from input signals representative of a set of images of the object taken from a plurality of respective camera positions, the apparatus comprising;

15 display means and control means operable to control the display means to display a model image derived from the model data and comprising a plurality of primitives for viewing by a user;

20 means for receiving at least one primitive selection signal responsive to user actuation of an input means whereby each primitive selection signal identifies a respective selected primitive of the model;



means for defining a plurality of virtual cameras in the three dimensional space having positions and look directions relative to the model which correspond substantially to those of the respective actual cameras relative to the object;

evaluating means for evaluating which of the virtual cameras is an optimum  
5 virtual camera for generating a view of the selected primitives; and

identifying means for identifying from the camera images a first camera image of the plurality of camera images taken from a camera position corresponding to that of the optimum virtual camera.

10 32. Apparatus as claimed in claim 31 comprising means for determining from the camera images a second camera image as being suitable for matching features in the first camera image, the control means being operable to control the display means to display the second camera image for comparison by the user with the first camera image.

15

33. Apparatus as claimed in claim 32 wherein the second camera image is taken from a camera position proximate to the optimum camera position.

34. Apparatus as claimed in any of claims 32 and 33 comprising means for  
20 receiving feature matching selection signals representative of user matched points in the first and second camera images.

35. Apparatus as claimed in claim 34 comprising means for generating updated model data to include additional detail corresponding to the received feature matching selection signals, means for rendering the updated model data to generate an updated model image and means for controlling the display means to display the updated  
5 model image.

36. Apparatus as claimed in any of claims 31 to 35 wherein the evaluating means comprises;

means for calculating for a selected primitive an aspect measurement  
10 representative of the visibility of the primitive when viewed in projection in the look direction of one of the virtual cameras;

means for repeating the calculating step to obtain a respective aspect measurement for each of the virtual cameras;

means for comparing the aspect measurements for the selected primitive and  
15 for determining a candidate virtual camera to be the virtual camera for which the corresponding aspect measurement is a maximum;

means for repeating the calculating, comparing and determining steps for each of the selected primitive whereby candidate virtual cameras are determined for each selected primitive; and

20 means for choosing the optimum virtual camera on the basis of the frequency with which virtual cameras are determined to be candidate virtual cameras.

37. Apparatus as claimed in claim 36 wherein the primitives comprise facets.

38. Apparatus as claimed in claim 37 wherein the means for calculation of the aspect measurement comprises, for a given facet and a given virtual camera, means  
5 for calculating a scalar product of a unit vector normal to the facet and a unit vector parallel to the look direction of the virtual camera.

39. Apparatus as claimed in claim 37 wherein the means for calculation of aspect measurement comprises means for calculating, for a given facet and for a given  
10 virtual camera, an area of the facet when viewed in projection in the look direction of the virtual camera.

40. Apparatus as claimed in any of claims 31 to 39 wherein the input means is a pointing means co-operable with the display means to provide input signals in the  
15 form of image co-ordinates of the displayed image.

41. Apparatus as claimed in any of claims 31 to 40 comprising means for generating the displayed model image by rendering the image data.

20 42. A storage medium storing processor implementable instructions for controlling a processor to carry out a method as claimed in any of claims 20 to 30.

43. An electrical signal carrying processor implementable instructions for controlling a processor to carry out a method as claimed in any of claims 20 to 30.

44. A computer program comprising processor implementable instructions for  
5 controlling a processor to carry out a method as claimed in any of claims 20 to 30.

45. In a method of operating an apparatus for processing image data in accordance with user selected co-ordinates of displayed images representative of said image data; an improvement wherein the apparatus performs the steps of;

10 displaying a first image representative of a first frame selected from said image data;

receiving pointing signals responsive to user actuation of a pointing device and displaying a cursor in the first image indicating an image point at a cursor position controlled by the pointing signals such that the cursor position is updated to track  
15 movement of the pointing device;

generating magnified image data representative of a first magnified image of a portion of the first image local to the cursor position and in fixed relationship thereto, and continuously updating the magnified image data in response to changes in the cursor position;

20 displaying the first magnified image simultaneously with the first image together with fiducial means indicating an image point in the first magnified image corresponding to the image point indicated in the first image at the cursor position;

and

receiving a selection signal responsive to user actuation of said pointing device and representative of co-ordinates of a first selected point in the first image indicated by the current cursor position.

5

46. In an apparatus for processing image data in accordance with user selected co-ordinates of displayed images representative of said image data; an improvement wherein the apparatus comprises;

display means operable to display a first image representative of a first frame  
10 selected from said image data;

pointing signal receiving means for receiving pointing signals responsive to user actuation of a pointing device and causing the display means to display a cursor in the first image indicating an image point at a cursor position controlled by the pointing signals such that the cursor position is updated to track movement of the  
15 pointing device;

generating means for generating magnified image data representative of a first magnified image of a portion of the first image local to the cursor position and in fixed relationship thereto, and for continuously updating the magnified image data in response to changes in the cursor position;

20 the display means being further operable to display the first magnified image simultaneously with the first image together with fiducial means indicating an image point in the first magnified image corresponding to the image point indicated in the

first image at the cursor position; and

selection signal receiving means for receiving a selection signal responsive to user actuation of said pointing device in use and representative of co-ordinates of a first selected point in the first image indicated by the current cursor position.

5

47. In an apparatus for processing image data in accordance with user selected co-ordinates of displayed images representative of said image data; a method wherein the apparatus performs the steps of;

10 displaying a first image representative of a first frame selected from said image data;

receiving pointing signals responsive to user actuation of a pointing device and displaying a cursor in the first image indicating an image point at a cursor position controlled by the pointing signals such that the cursor position is updated to track movement of the pointing device;

15 generating magnified image data representative of a first magnified image of a portion of the first image local to the cursor position and in fixed relationship thereto, and continuously updating the magnified image data in response to changes in the cursor position;

20 displaying the first magnified image simultaneously with the first image together with fiducial means indicating an image point in the first magnified image corresponding to the image point indicated in the first image at the cursor position; and

receiving a selection signal responsive to user actuation of said pointing device and representative of co-ordinates of a first selected point in the first image indicated by the current cursor position.

5 48. In a method of operating an apparatus for generating model data representative of a model in a three dimensional space of an object from input signals representative of a set of images of the object taken from a plurality of respective camera positions, an improvement wherein the apparatus performs the steps of;

10 displaying a model image derived from the model data and comprising a plurality of primitives for viewing by a user;

receiving at least one primitive selection signal responsive to user actuation of an input means whereby each primitive selection signal identifies a respective selected primitive of the model;

15 defining a plurality of virtual cameras in the three dimensional space having positions and look directions relative to the model which correspond substantially to those of the respective actual cameras relative to the object;

evaluating which of the virtual cameras is an optimum virtual camera for generating a view of the selected primitives;

20 identifying from the camera images a first camera image of the plurality of camera images taken from a camera position corresponding to that of the optimum virtual camera.

49. In an apparatus for generating model data representative of a model in a three dimensional space of an object from input signals representative of a set of images of the object taken from a plurality of respective camera positions, an improvement whereby the apparatus comprises;

5 display means and control means operable to control the display means to display a model image derived from the model data and comprising a plurality of primitives for viewing by a user;

means for receiving at least one primitive selection signal responsive to user actuation of an input means whereby each primitive selection signal identifies a  
10 respective selected primitive of the model;

means for defining a plurality of virtual cameras in the three dimensional space having positions and look directions relative to the model which correspond substantially to those of the respective actual cameras relative to the object;

evaluating means for evaluating which of the virtual cameras is an optimum  
15 virtual camera for generating a view of the selected primitives; and

identifying means for identifying from the camera images a first camera image of the plurality of camera images taken from a camera position corresponding to that of the optimum virtual camera.

20 50. In an apparatus for generating model data representative of a model in a three dimensional space of an object from input signals representative of a set of images of the object taken from a plurality of respective camera positions, an improvement



whereby the apparatus performs the steps of;

displaying a model image derived from the model data and comprising a plurality of primitives for viewing by a user;

receiving at least one primitive selection signal responsive to user actuation  
5 of an input means whereby each primitive selection signal identifies a respective selected primitive of the model;

defining a plurality of virtual cameras in the three dimensional space having positions and look directions relative to the model which correspond substantially to those of the respective actual cameras relative to the object;

10 evaluating which of the virtual cameras is an optimum virtual camera for generating a view of the selected primitives;

identifying from the camera images a first camera image of the plurality of camera images taken from a camera position corresponding to that of the optimum virtual camera.

ABSTRACTIMAGE PROCESSING APPARATUS

A method of processing image data entails the input of user selected image co-ordinates using a display interface in which a mouse is used to control the position of  
5 a cursor. A magnified image is additionally displayed in a magnified image window which overlays a corner portion of the main image window. The magnified image includes a fixed graticule and displays a magnified portion of the image which tracks the current position of the cursor, thereby enabling the user to locate the cursor with greater accuracy with reference to the magnified image. After selection of an image  
10 point, the image displayed in the magnified window is frozen. Matching image point co-ordinates in two separate images may be input to the processor by providing respective magnified image windows in each of the main image windows. The magnified images thereby provided enable pairs of matching co-ordinates to be rapidly selected without the need for the user to input any additional instructions to control  
15 the generation of the magnified images. The method is particularly useful in providing matching image co-ordinates for generating a three-dimensional model based on a series of image frames representative of different views of the object. A further aspect of the invention provides for a pair of camera images to be automatically selected in response to the user selecting a set of primitives in the model image.

20 (Figure 3 refers)

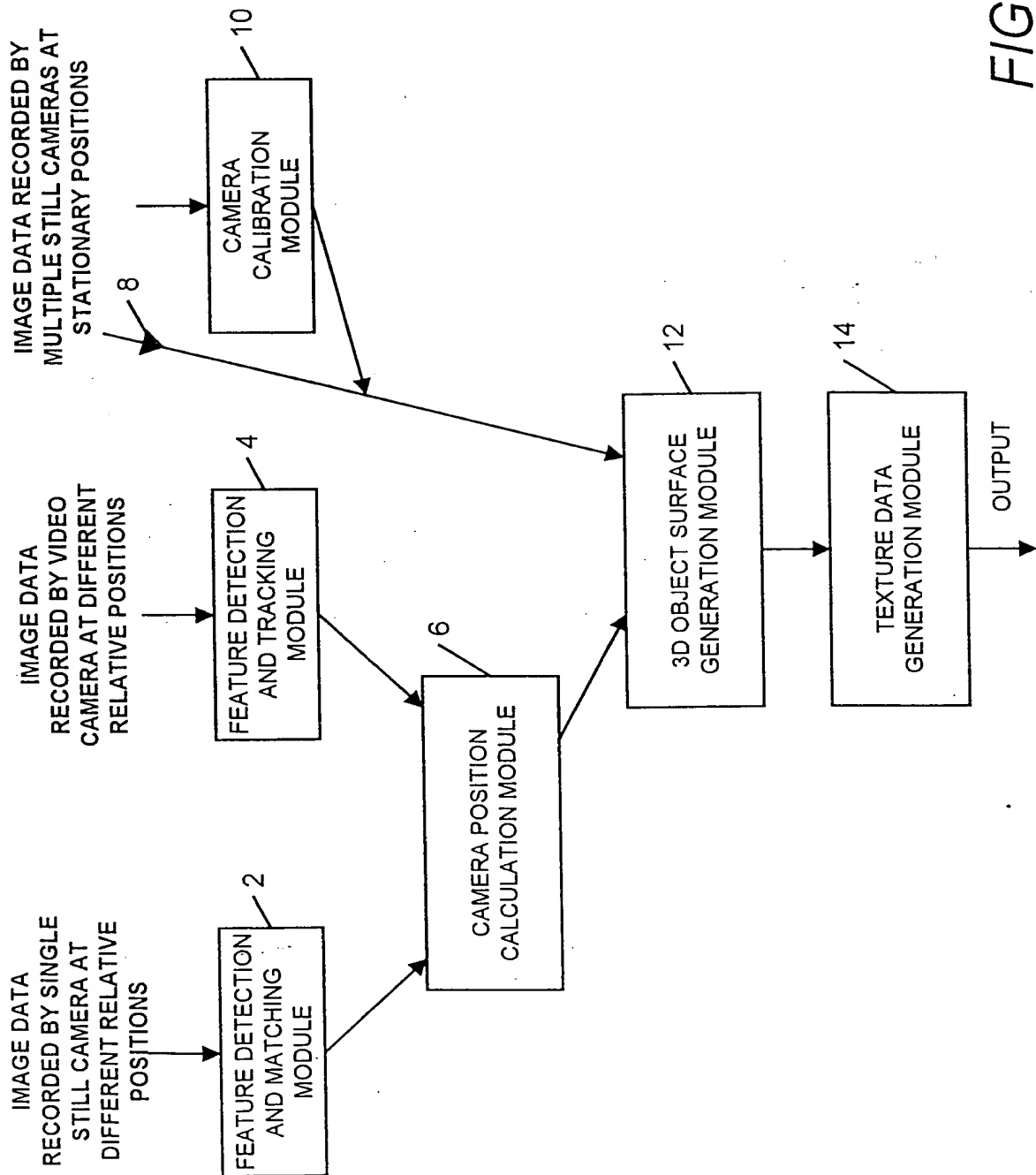
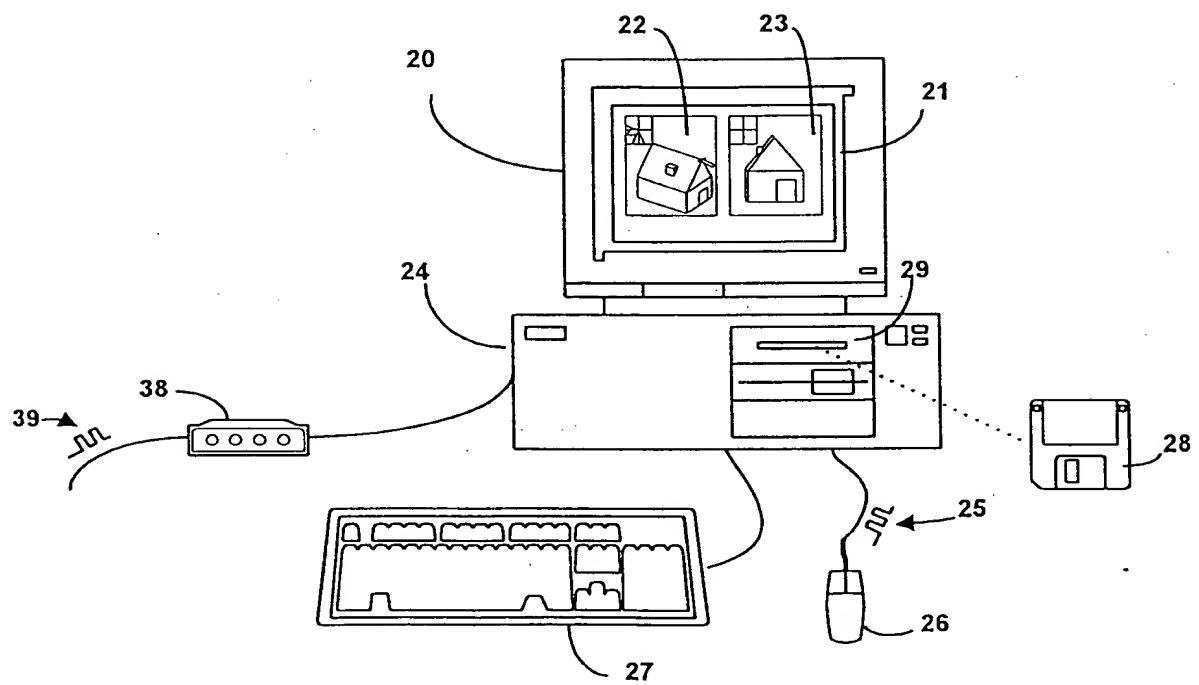


FIG. 1

**THIS PAGE BLANK (USPTO)**

Fig 2



**THIS PAGE BLANK (USPTO)**

Fig 3

FIRST PHASE

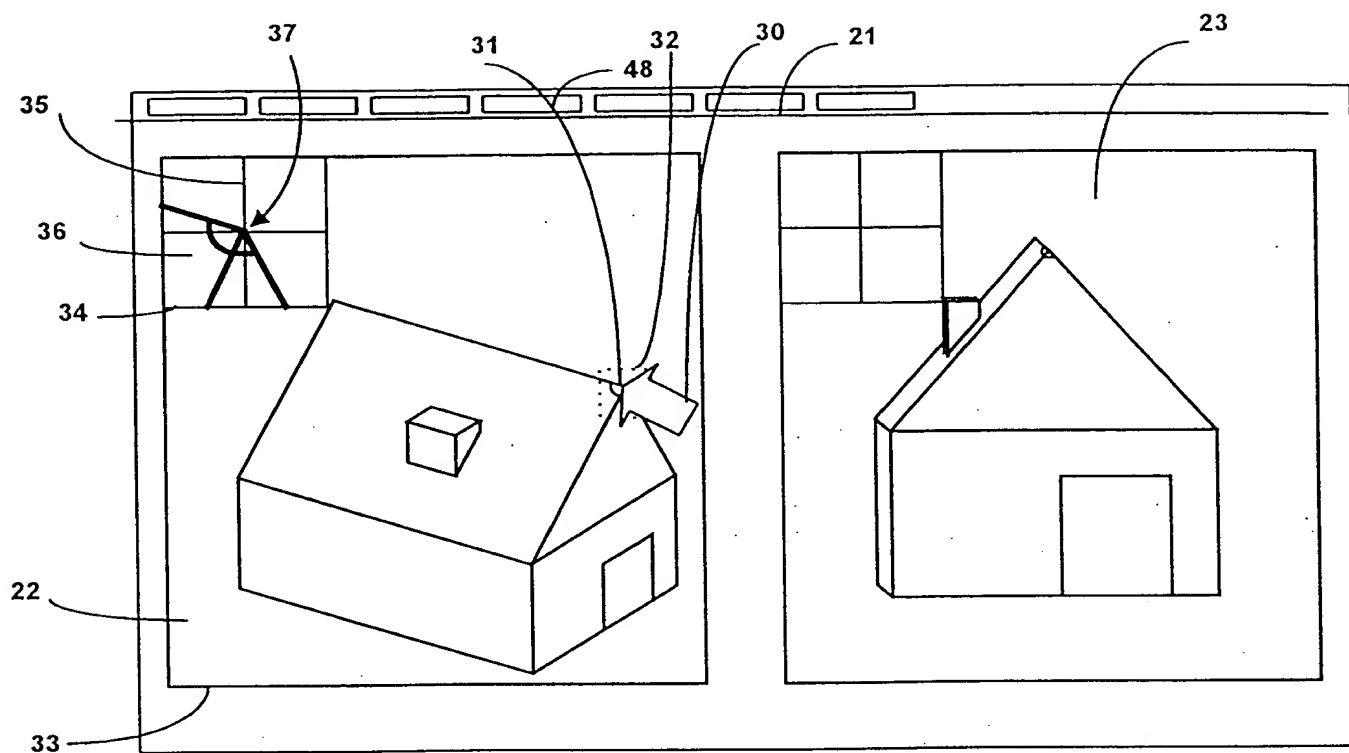
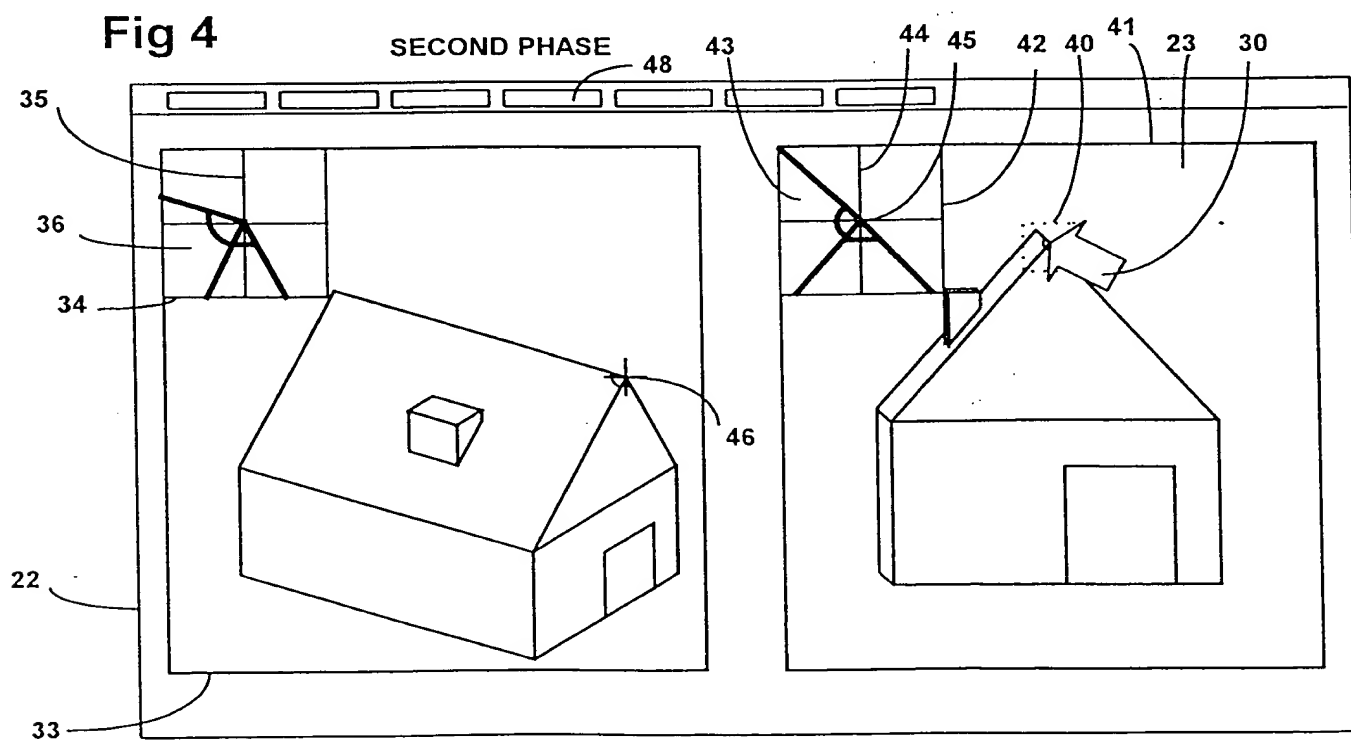


Fig 4

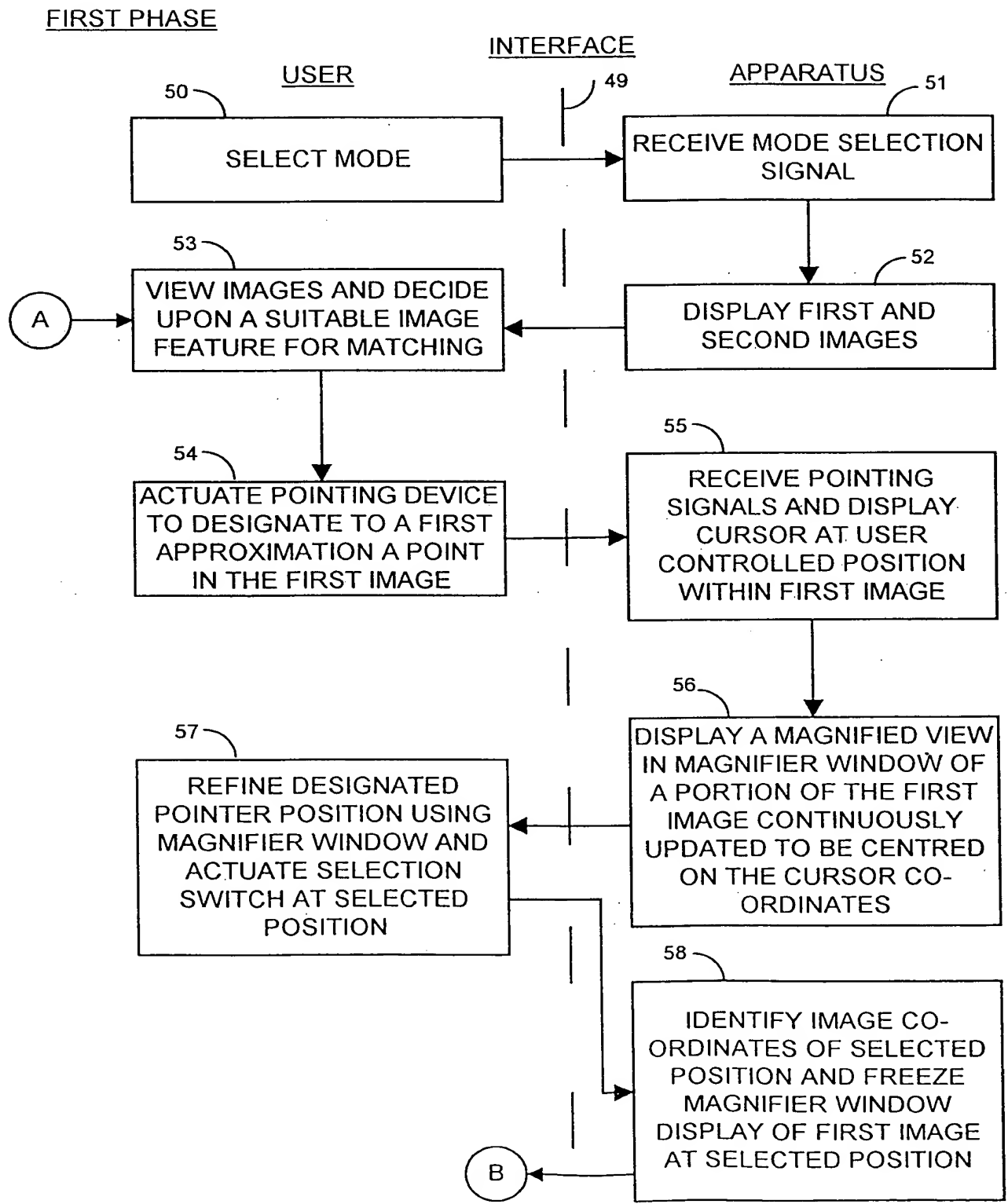
SECOND PHASE



**THIS PAGE BLANK (USPTO)**



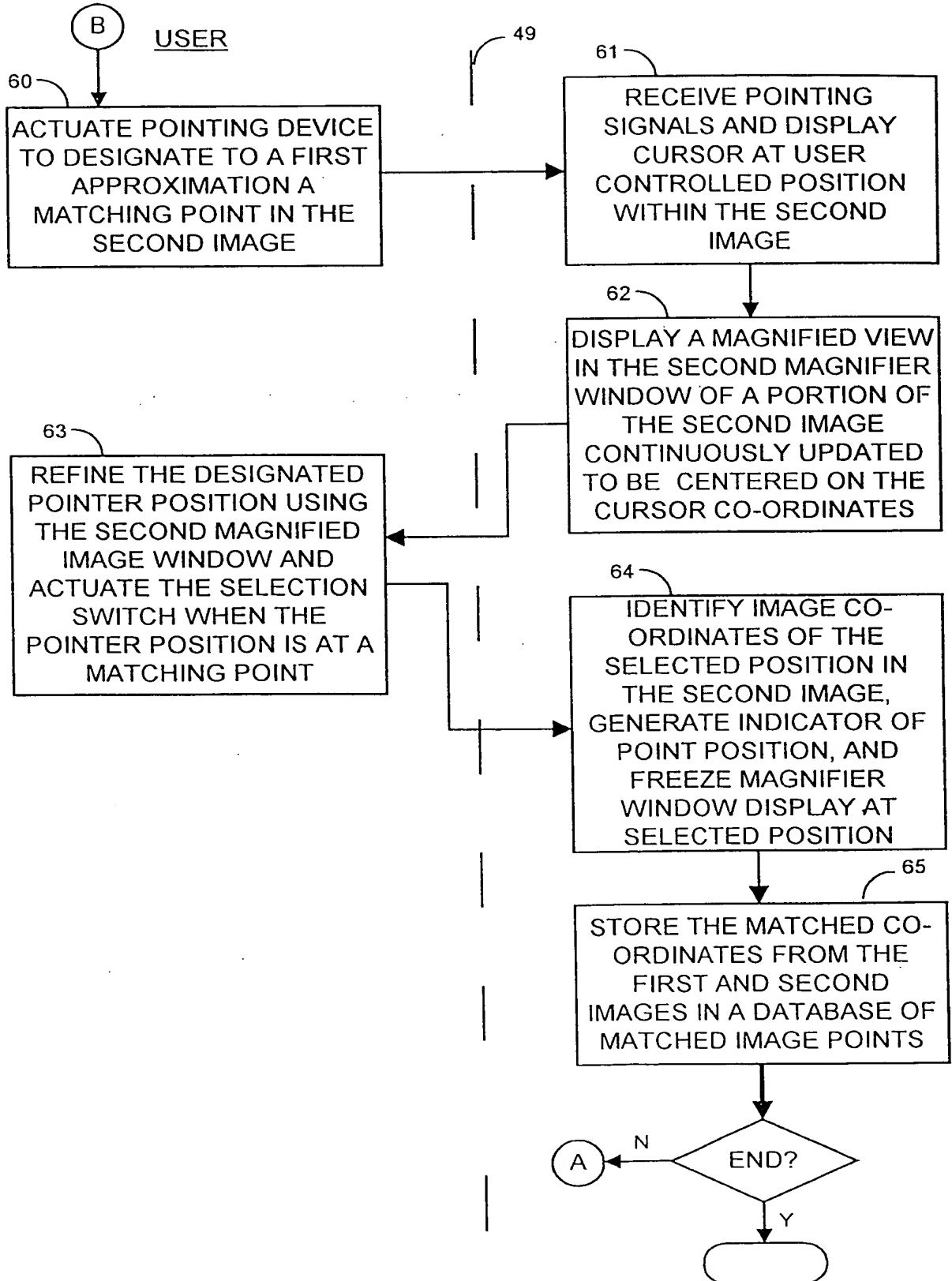
FIG. 5



**THIS PAGE BLANK (USPTO)**

5/23

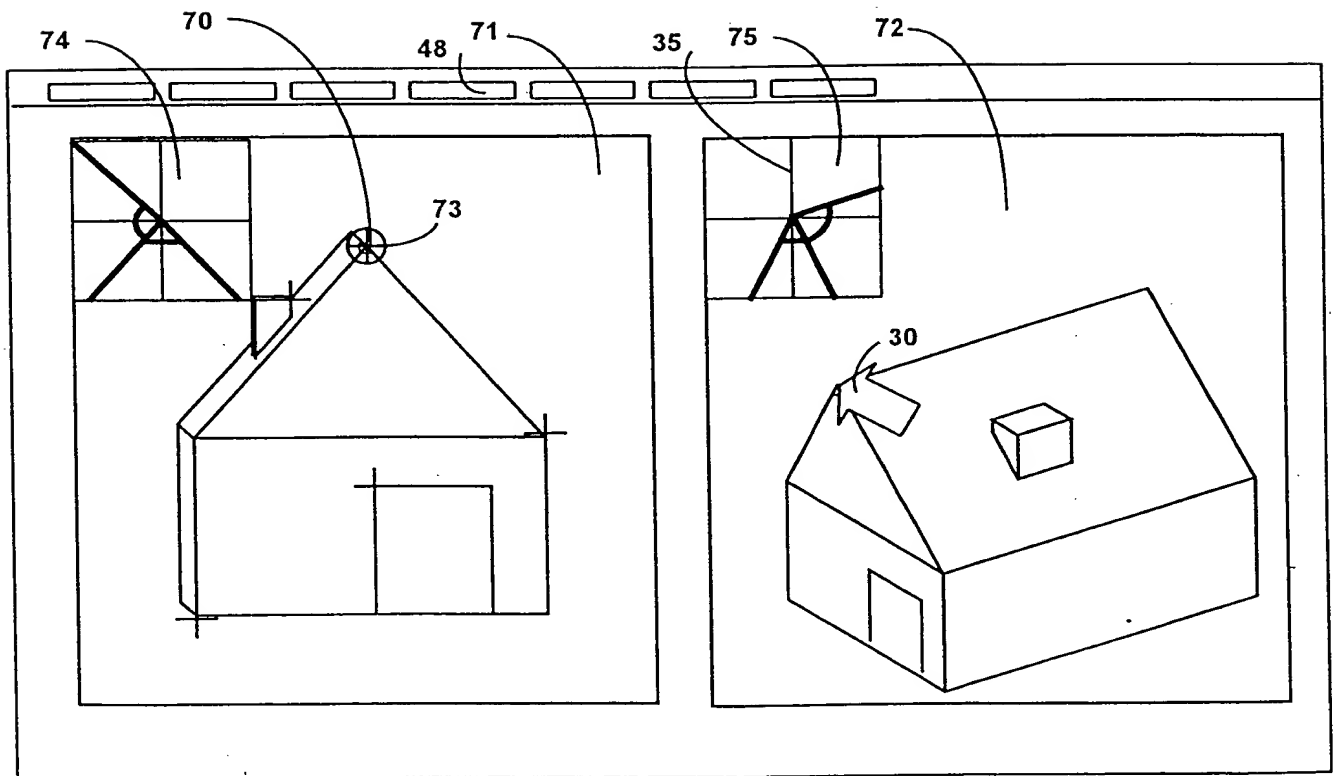
FIG. 6

SECOND PHASEINTERFACEAPPARATUS

**THIS PAGE BLANK (USPTO)**

6/23

Fig 7

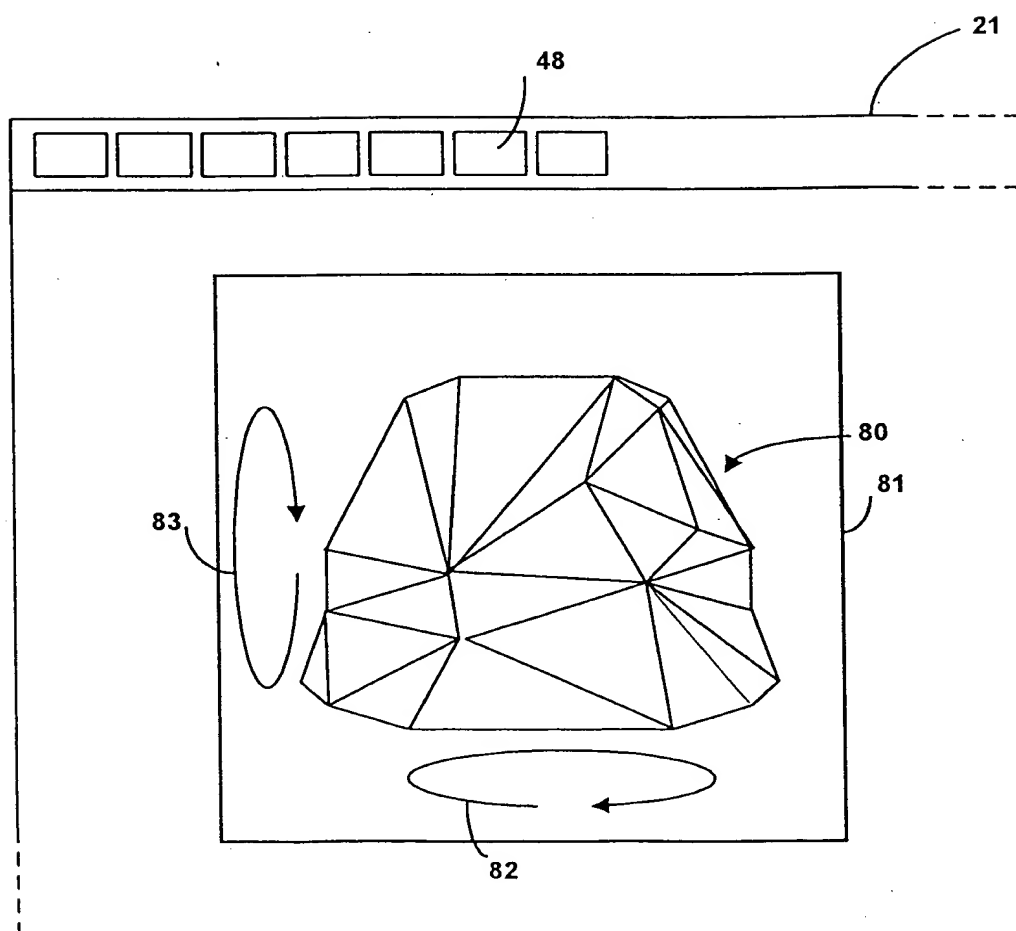


**THIS PAGE BLANK (USPTO)**

7/23

Fig 8

INITIAL ORIENTATION OF MODEL IMAGE

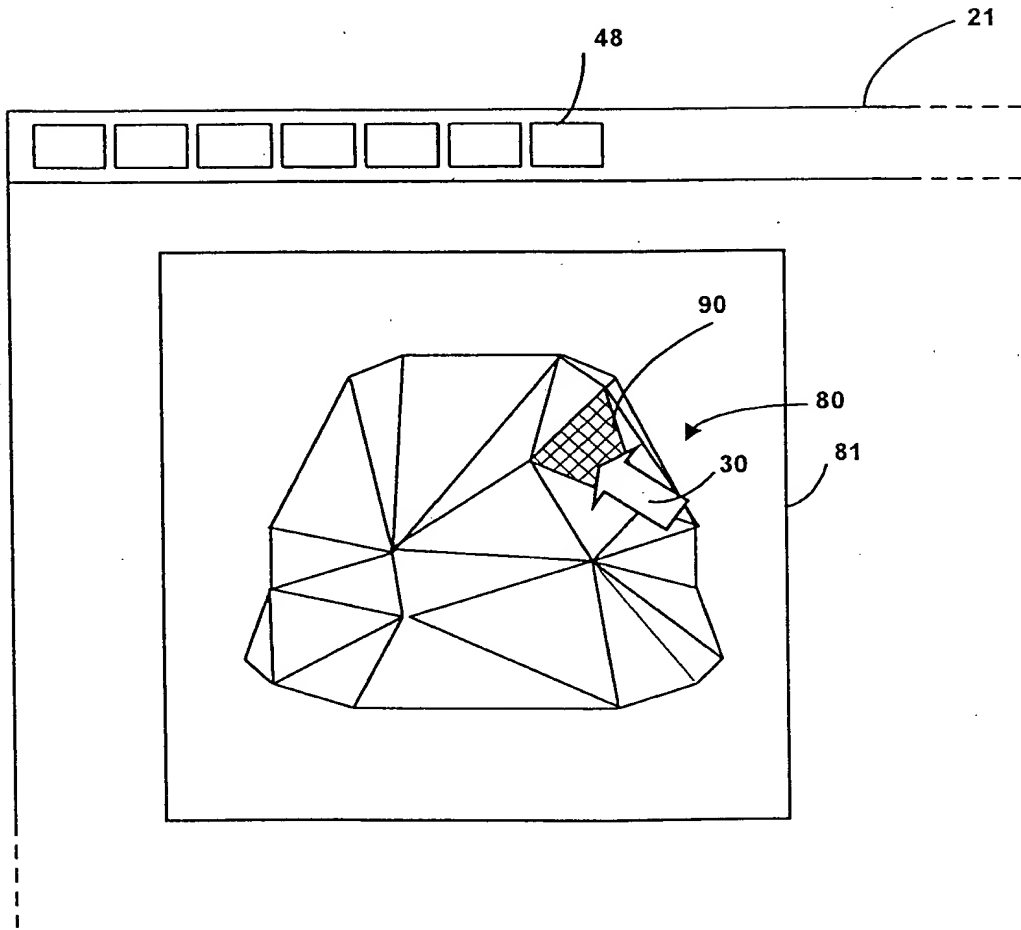


**THIS PAGE BLANK (USPTO)**



Fig 9

FACET SELECTION IN MODEL IMAGE

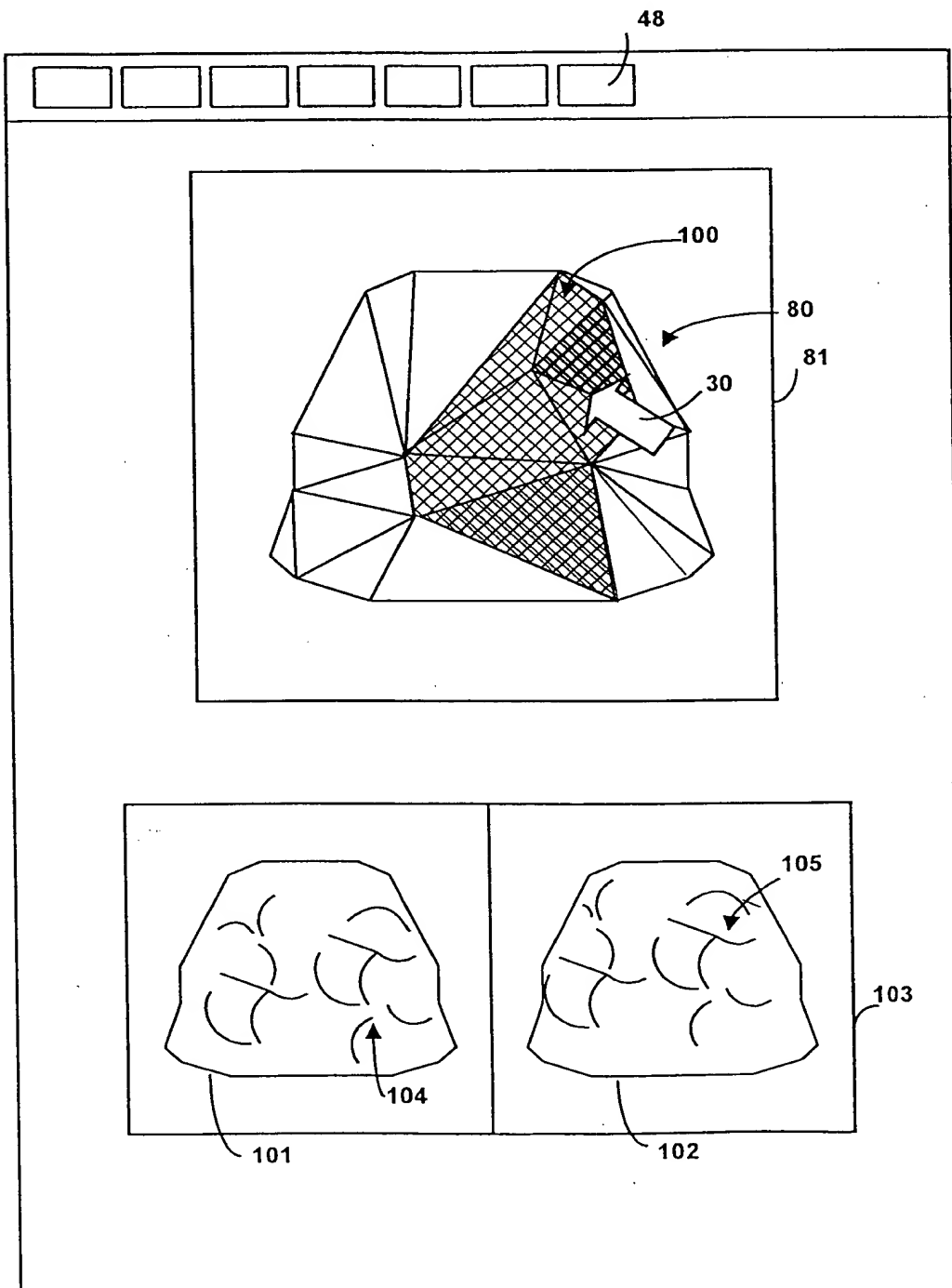


**THIS PAGE BLANK (USPTO)**

9/23

**Fig 10**

SELECTION OF MULTIPLE FACETS AND DISPLAY OF OPTIMUM AND  
COMPLEMENTARY CAMERA IMAGES

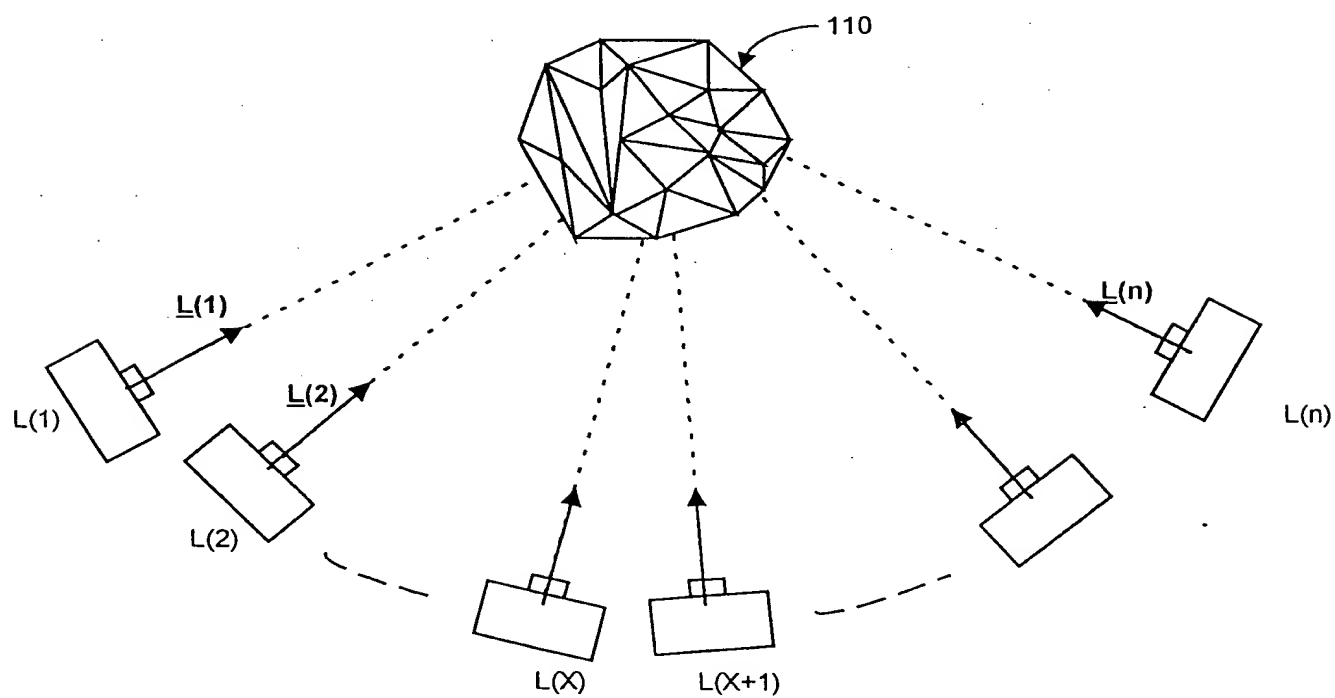


**THIS PAGE BLANK (USPTO)**

10/23

**Fig 11**

**VIRTUAL CAMERA POSITIONS RELATIVE TO THE MODEL**

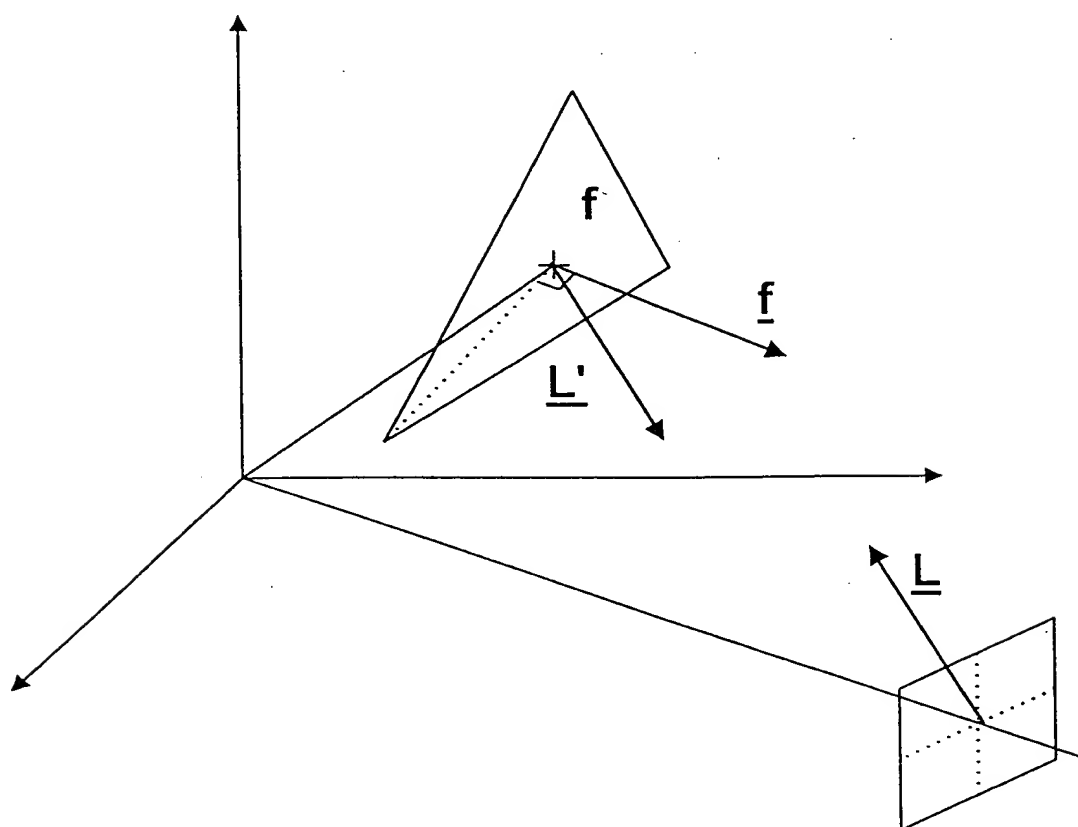


**THIS PAGE BLANK (USPTO)**

11/23

Fig 12

UNIT VECTORS FOR ASPECT MEASUREMENT



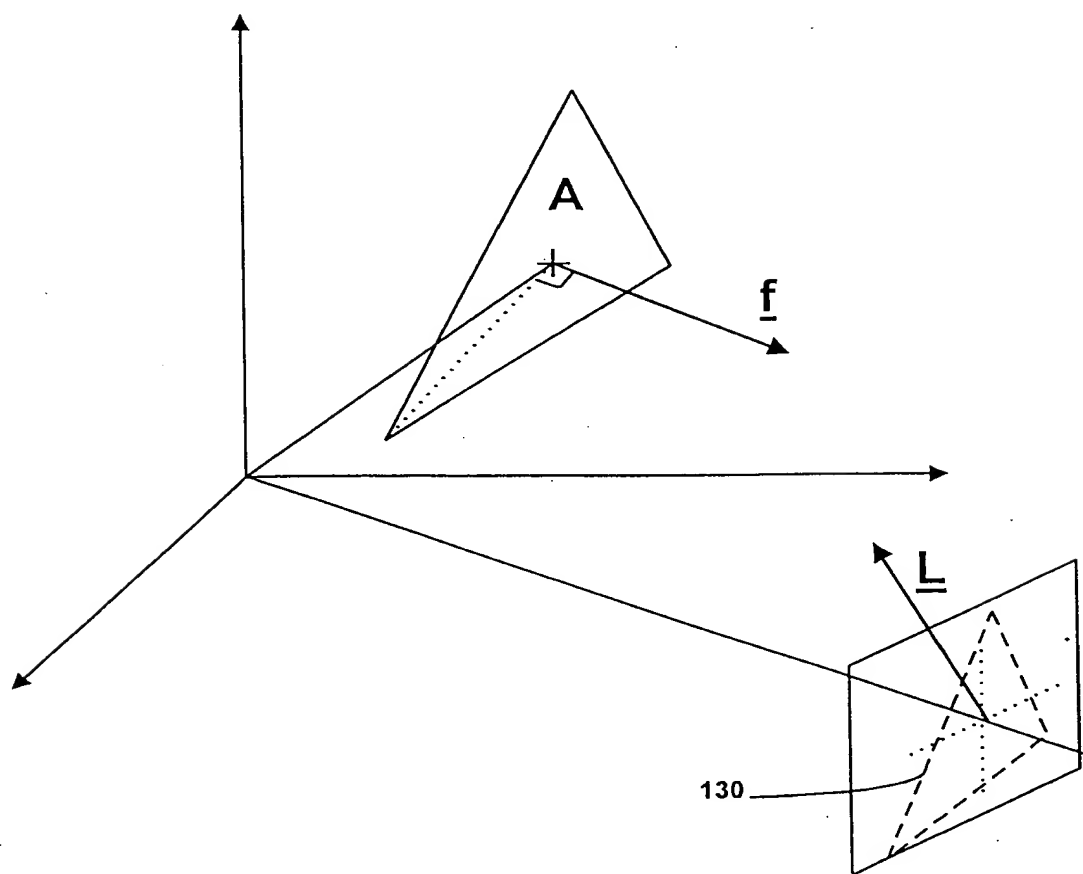
**THIS PAGE BLANK (USPTO)**



12/23

**Fig 13**

VISIBLE AREA MEASUREMENT FOR A GIVEN FACET

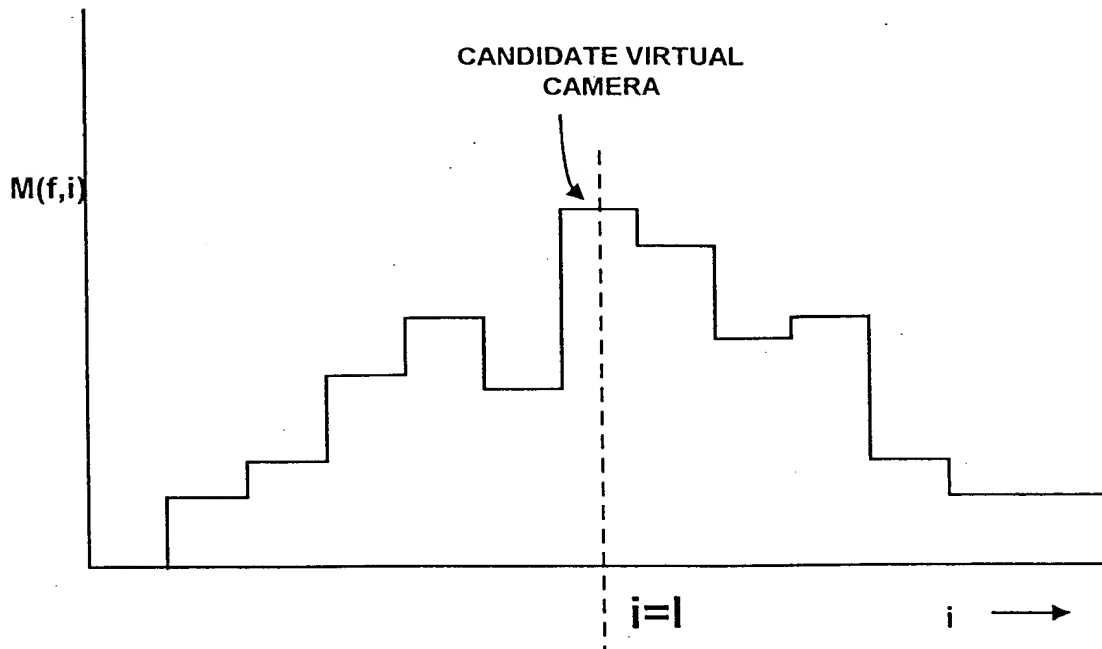


**THIS PAGE BLANK (USPTO)**

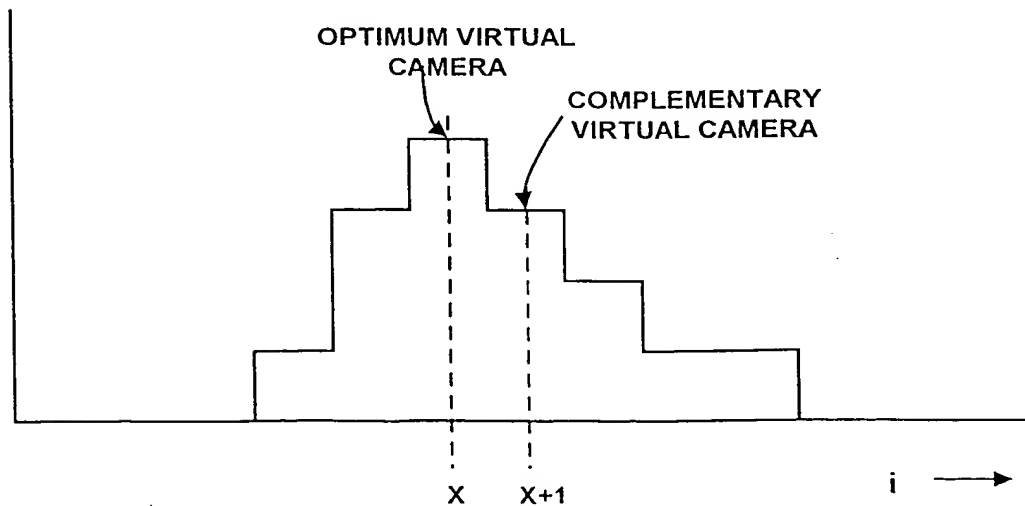
13/23

**Fig 14**

HISTOGRAM OF ASPECT MEASUREMENT FOR A GIVEN FACET  
FOR EACH VIRTUAL CAMERA

**Fig 15**

HISTOGRAM SHOWING FREQUENCY WITH WHICH VIRTUAL  
CAMERAS ARE SELECTED AS CANDIDATE CAMERAS

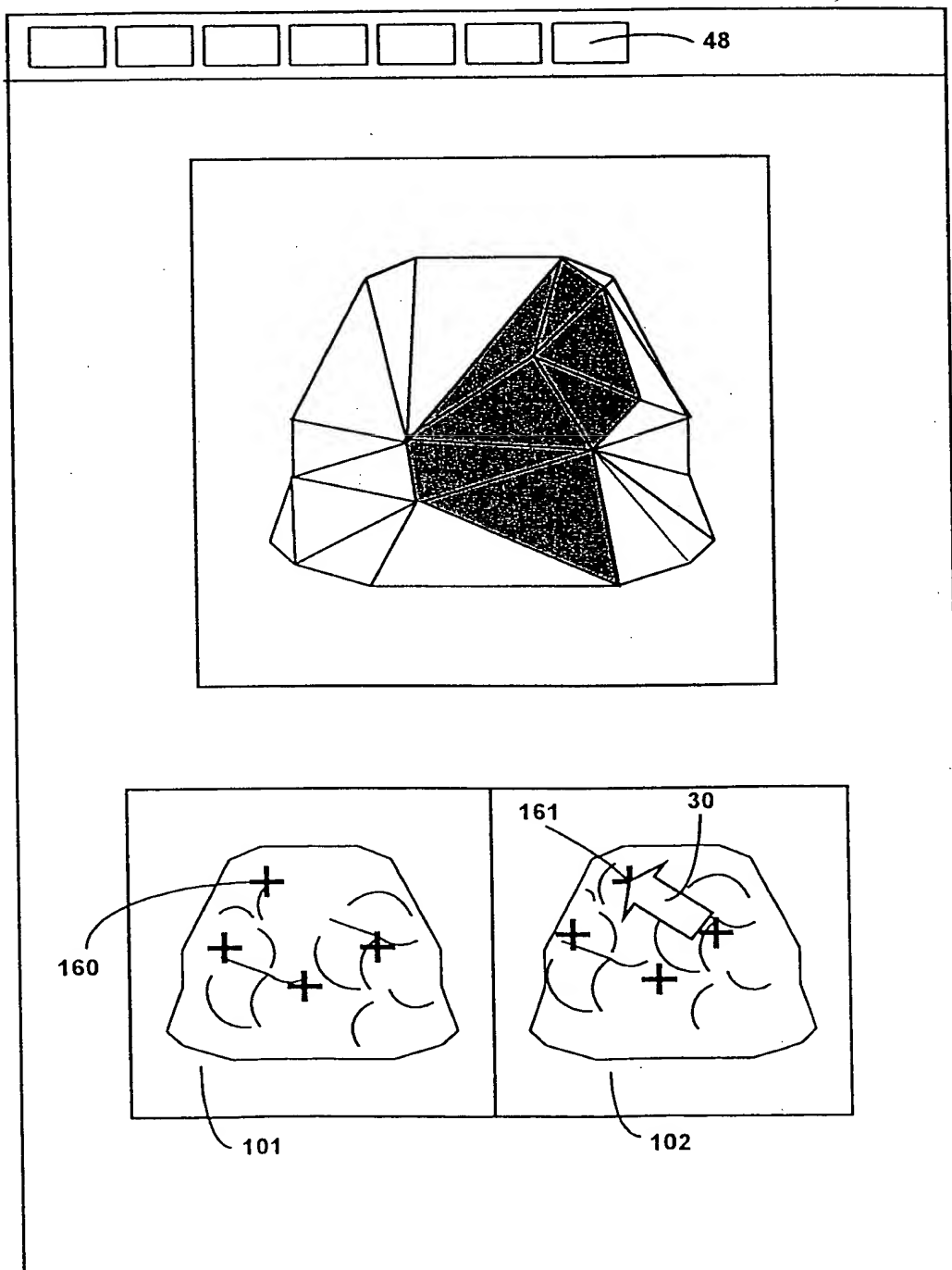


**THIS PAGE BLANK (USPTO)**

14/23

Fig 16

UPDATING MODEL BY MATCHING POINTS

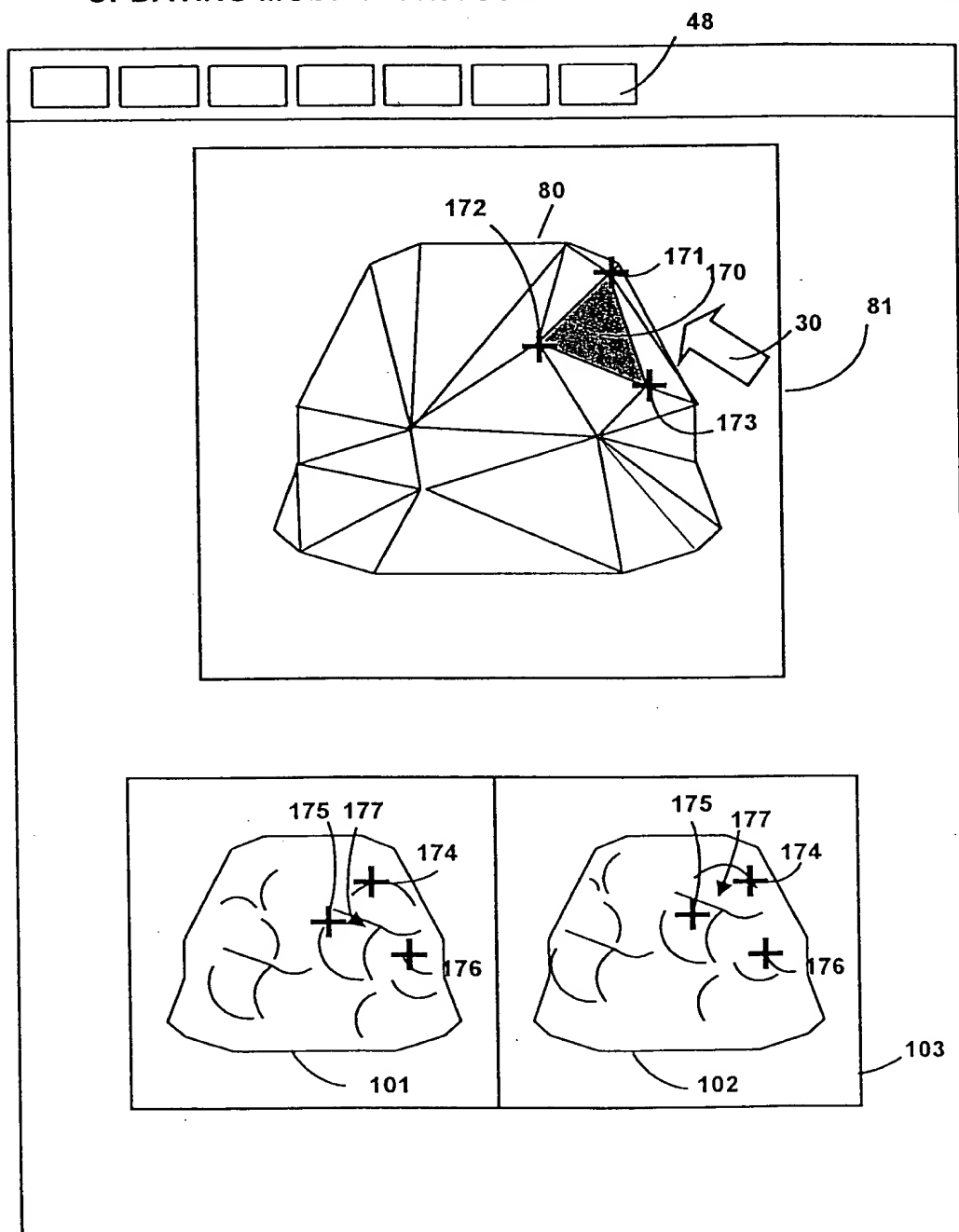


**THIS PAGE BLANK (USPTO)**

15/23

Fig 17A

UPDATING MODEL DATA USING DRAG AND DROP METHOD



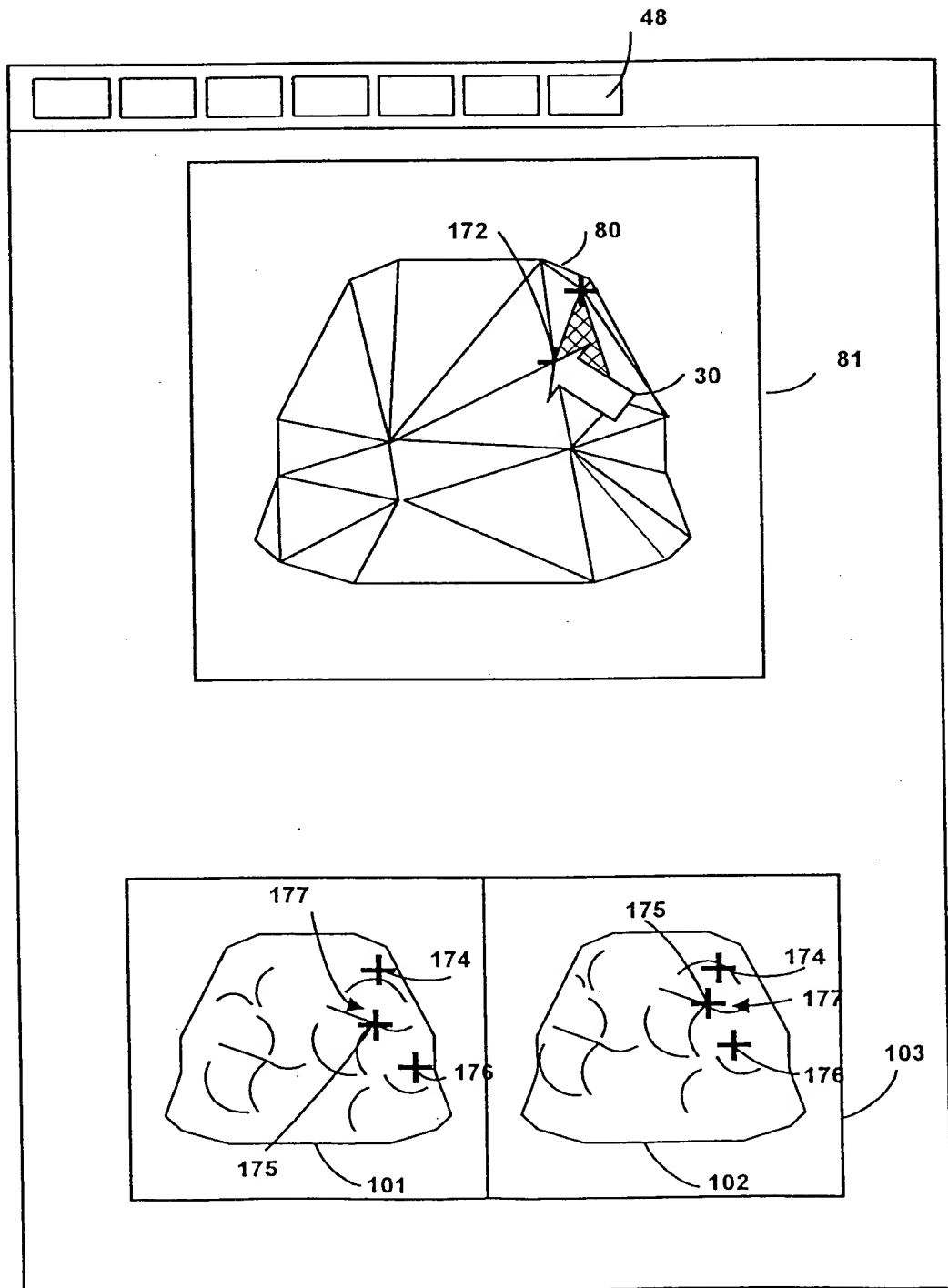
**THIS PAGE BLANK (USPTO)**



16/23

**Fig 17B**

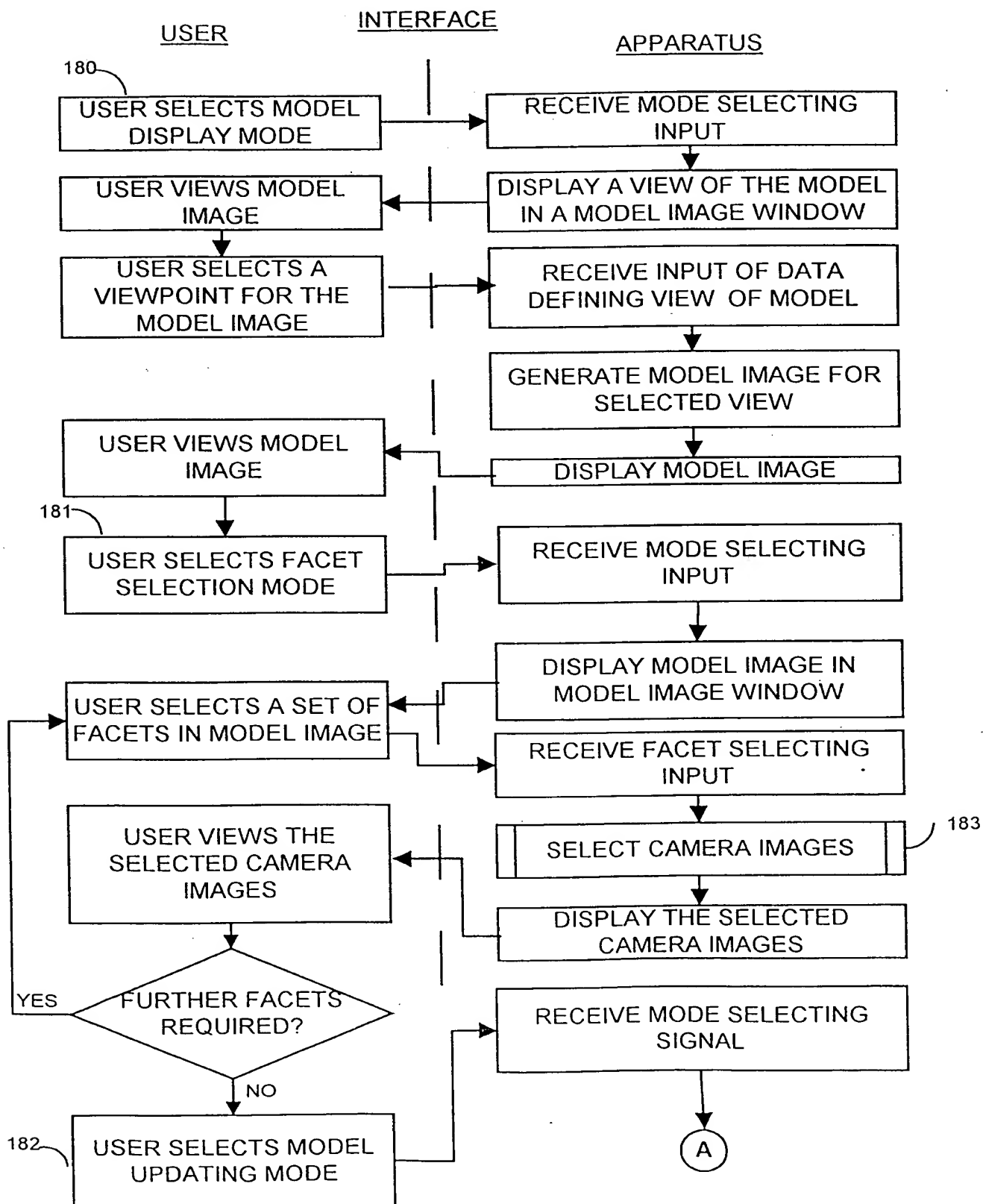
MODEL POINT MOVED BY USING DRAG AND DROP METHOD



**THIS PAGE BLANK (USPTO)**

17/23

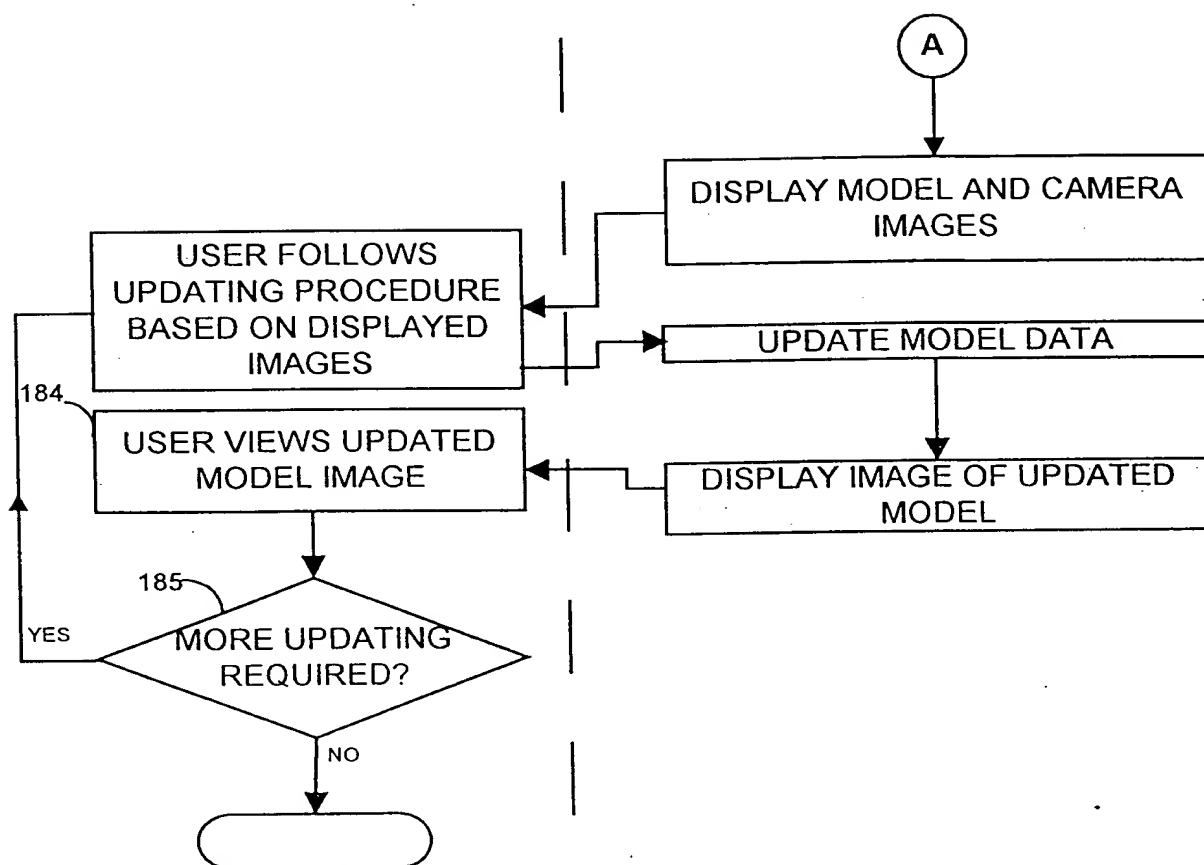
FIG. 18A



**THIS PAGE BLANK (USPTO)**

18/23

FIG. 18B  
INTERFACE

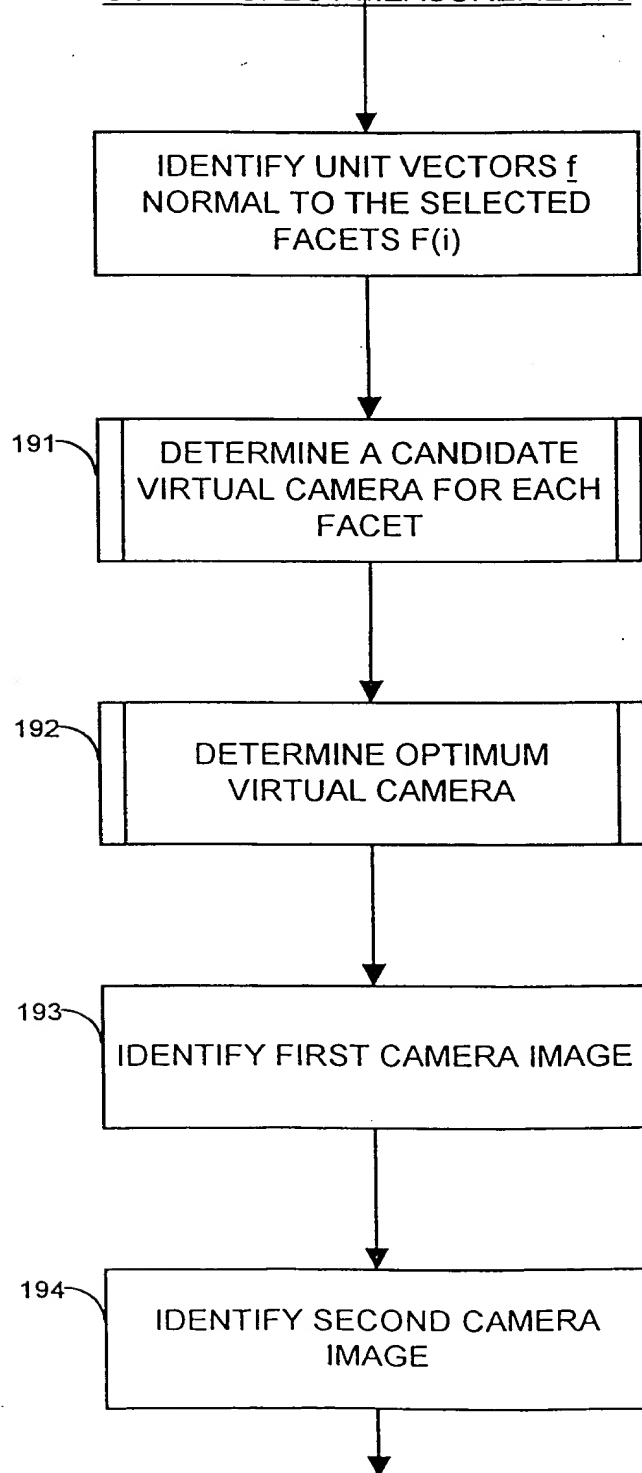


**THIS PAGE BLANK (USPTO)**

19/23

FIG. 19

SELECTION OF OPTIMISED CAMERA IMAGES  
USING ASPECT MEASUREMENTS



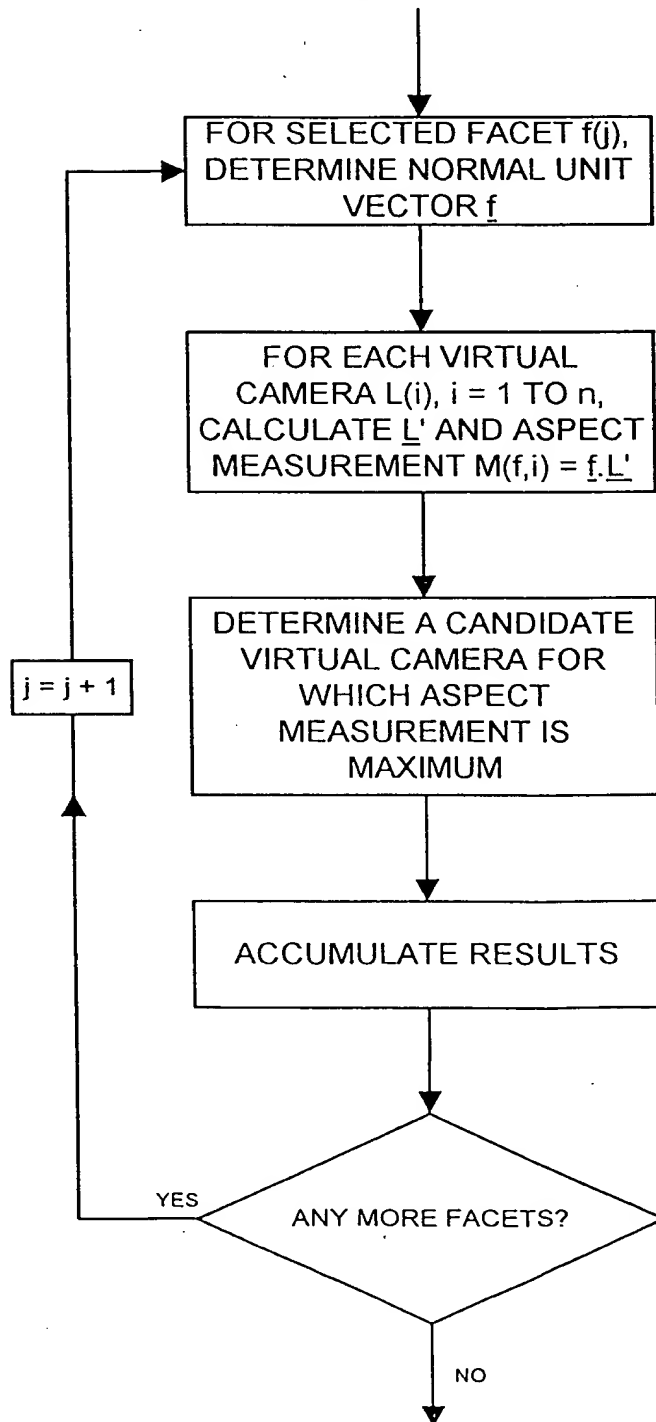
**THIS PAGE BLANK (USPTO)**



20/23

FIG. 20

DETERMINATION OF CANDIDATE VIRTUAL CAMERA  
FOR EACH FACET BASED ON ASPECT  
MEASUREMENTS

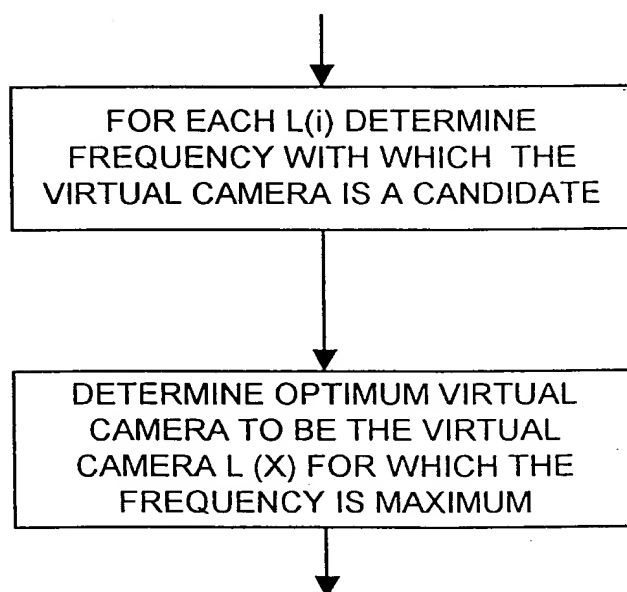


**THIS PAGE BLANK (USPTO)**

21/23

FIG. 21

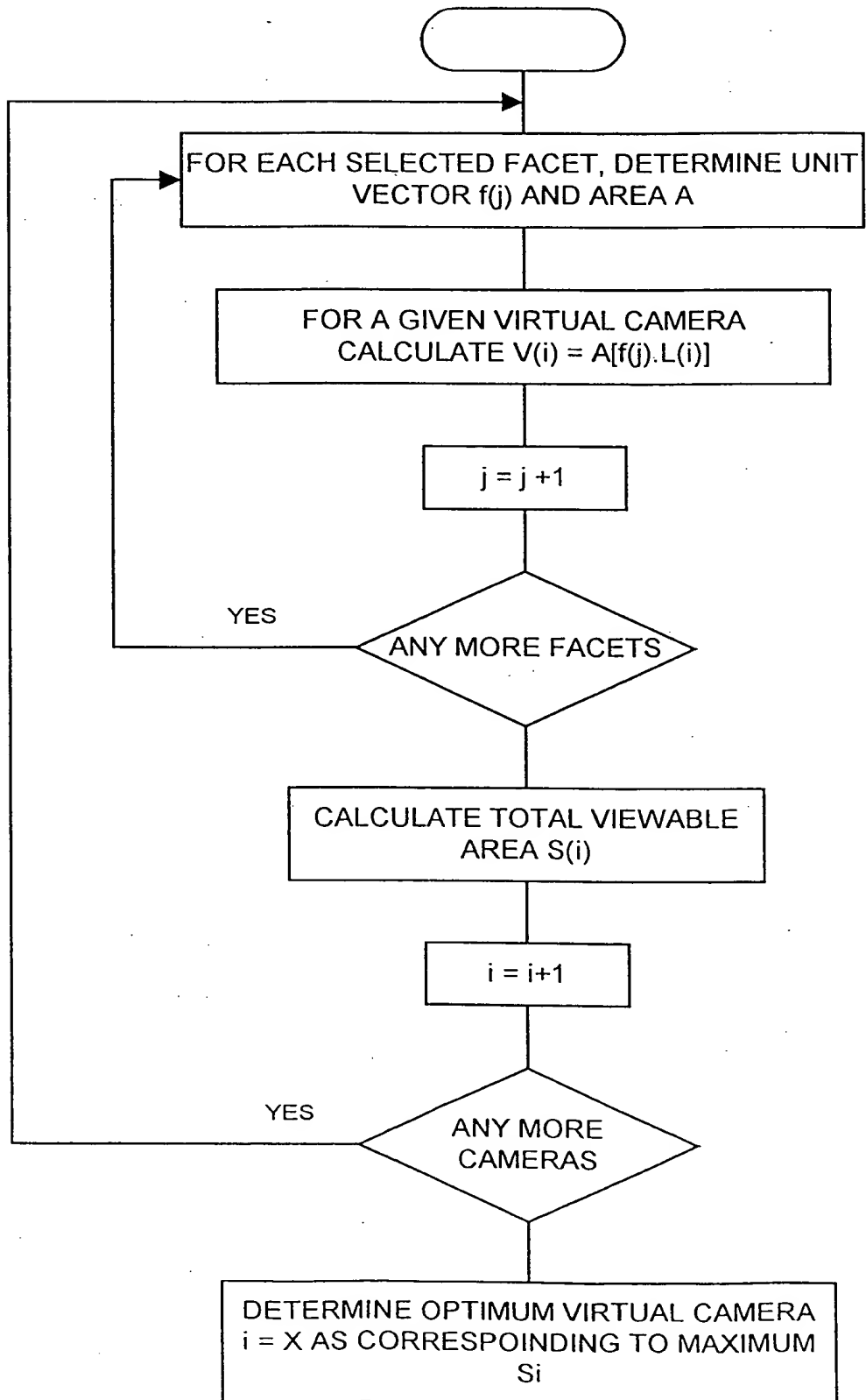
DETERMINATION OF OPTIMUM VIRTUAL CAMERA BASED  
ON ASPECT MEASUREMENTS



**THIS PAGE BLANK (USPTO)**

FIG. 22

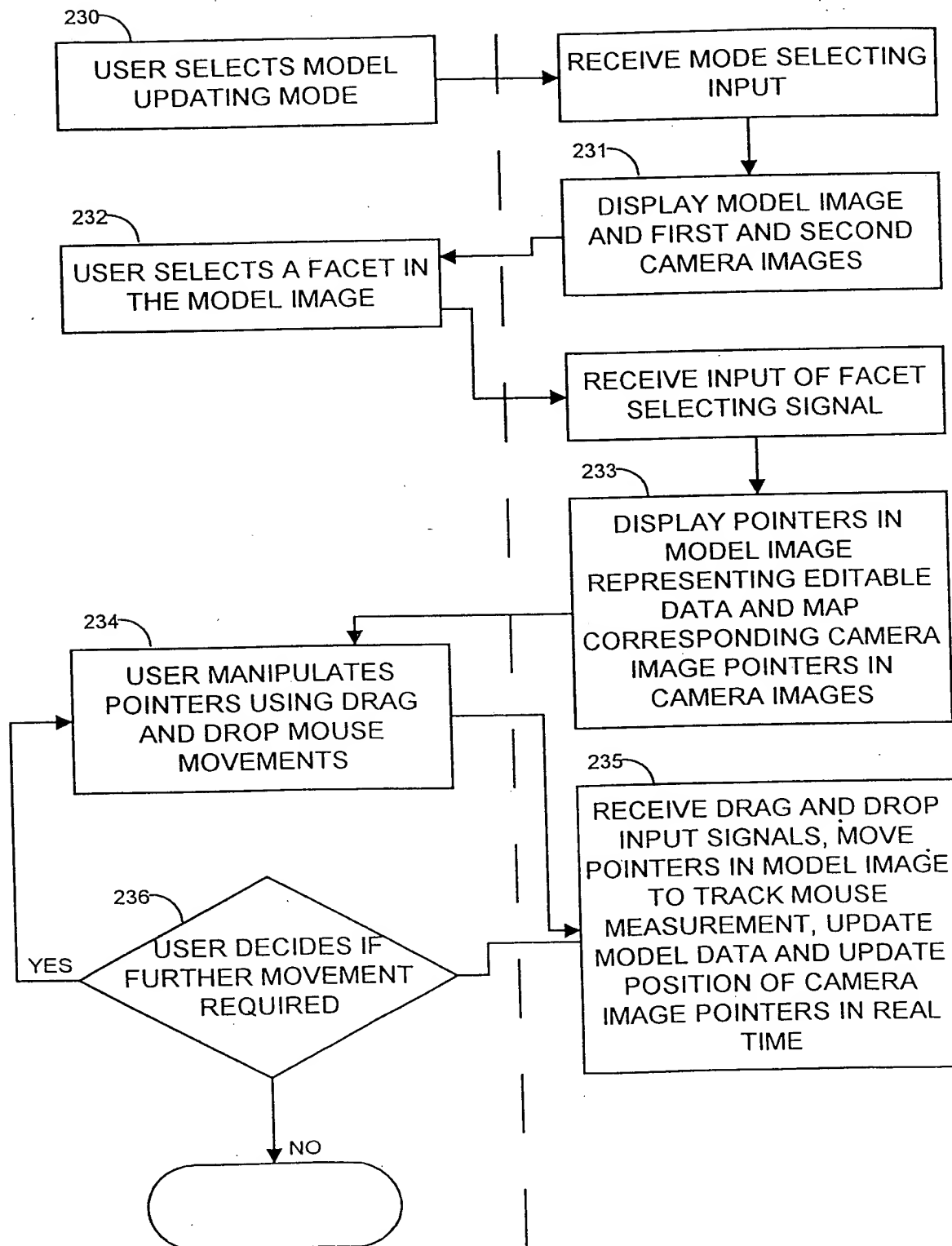
DETERMINATION OF OPTIMUM VIRTUAL CAMERA BASED  
ON VIEWABLE AREA MEASUREMENTS



**THIS PAGE BLANK (USPTO)**

23/23

FIG. 23



**THIS PAGE BLANK (USPTO)**